STRUCTURE-SPECIFIC PROBABILISTIC SEISMIC RISK ASSESSMENT

A THESIS
SUBMITTED TO THE UNIVERSITY OF CANTERBURY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Brendon Archie Bradley
April 2009

Department of Civil and Natural Resources Engineering
College of Engineering
University of Canterbury
Christchurch, New Zealand

Thesis Abstract

This thesis addresses a diverse range of topics in the area of probabilistic seismic risk analysis of engineering facilities. This intentional path of diversity has been followed primarily because of the relatively new and rapid development of this facet of earthquake engineering. As such this thesis focuses on the rigorous scrutinization of current, and in particular, simplified methods of seismic risk assessment; the development of novel aspects of a risk assessment methodology which provides easily communicated performance measures and explicit consideration for the many uncertainties in the entire earthquake problem; and the application of this methodology to case-study examples including structures supported on pile foundations embedded in liquefiable soils.

The state-of-the-art in seismic risk and loss assessment is discussed via the case study of a 10 storey New Zealand office building. Particular attention is given to the quality and quantity of information that such assessment methodologies provide to engineers and stakeholders for rational decision-making.

Two chapters are devoted to the investigation of the power-law model for representing the ground motion hazard. Based on the inaccuracy of the power-law model at representing the seismic hazard over a wide range of exceedance rates, an alternative, more accurate, parametric hazard model based on a hyperbola in log-log space is developed and applied to New Zealand peak ground acceleration and spectral acceleration hazard data. A semi-analytical closed-form solution for the demand hazard is also developed using the hyperbolic hazard model and applied for a case-study performance assessment. The power-law hazard model is also commonly used to obtain a closed-form solution for the annual rate of structural collapse (collapse hazard). The magnitude of the error in this closed-form solution due to errors in the necessary functional forms of its constitutive relations is examined via a parametric study.

A series of seven chapters are devoted to the further development of various aspects of a seismic risk assessment methodology. Intensity measures for use in the estimation of spatially distributed seismic demands and seismic risk assessment which are: easily predicted; can predict seismic response with little uncertainty; and are unbiased regarding

additional properties of the input ground motions are examined. An efficient numerical integration algorithm which is specifically tailored for the solution of the governing risk assessment equations is developed and compared against other common methods of numerical integration. The efficacy of approximate uncertainty propagation in seismic risk assessment using the so-called First-Order Second-Moment method is investigated. Particular attention is given to the locations at which the approximate uncertainty propagation is used, the possible errors for various computed seismic risk measures, and the reductions in computational demands. Component correlations have to date been not rigorously considered in seismic loss assessments due to complications in their estimation and tractable methodologies to account for them. Rigorous and computationally efficient algorithms to account for component correlations are presented. Particular attention is also given to the determination of correlations in the case of limited empirical data, and the errors which may occur in seismic loss assessment computations neglecting proper treatment of correlations are examined. Trends in magnitude, distribution, and correlation of epistemic uncertainties in seismic hazard analyses for sites in the San Francisco bay area are examined. The characteristics of these epistemic uncertainties are then used to compare and contrast three methods which can be used to propagate such uncertainties to other seismic risk measures. Causes of epistemic uncertainties in component fragility functions, their evaluation, and combination are also examined.

A series of three chapters address details regarding the seismic risk assessment of structures supported on pile foundations embedded in liquefiable soils. A ground motion prediction equation for spectrum intensity (found to be a desirable intensity measure for seismic response analysis in liquefiable soils) is developed based on ground motion prediction equations for spectral accelerations, which are available in abundance in literature. Determination of intensity measures for the seismic response of pile foundations, which are invariably located in soil deposits susceptible to liquefaction, is examined. Finally, a rigorous seismic performance and loss assessment of a case-study bridge structure is examined using rigorous ground motion selection, seismic effective stress analyses, and professional cost estimates. Both direct repair and loss of functionality consequences for the bridge structure are examined.

Acknowledgements

Firstly I would like to thank my supervisory team of Dr. Rajesh Dhakal, Dr. Misko Cubrinovski, Dr. Gregory MacRae, and Dr. Dominic Lee. This thesis has involved the combination of many facets of earthquake engineering and each of you has provided a significant and unique contribution which I am greatly indebted for. As well as your guidance in the technical aspects of this research you also have had a positive impact in my personal development for which I am further grateful to you all.

I would also like to acknowledge significant financial support provided to me by the Tertiary Education Commission Bright Futures Scheme, New Zealand Freemasons Society, Fulbright New Zealand, New Zealand Society for Earthquake Engineering, New Zealand Earthquake Commission, and the University of Canterbury. Without their support, this thesis and the variety of opportunities I have experienced would not be possible.

This thesis and my professional development have also greatly benefited from interactions with many individuals. To my fellow postgraduate students, Mr. Weng Yuen Kam, Mr. Brian Peng, Mr. Alistar Boys, Mr. Perry Jackson and numerous others, thank you for your providing insightful comments and suggestions on a range of topics, as well as making being a postgraduate student several very enjoyable years. To other people in New Zealand and abroad: Dr. Athol Carr, Mr. Hayden Bowen, Dr. Graeme McVerry, Dr. Mark Stirling, Dr. Jack Baker, Dr. Nico Luco, Dr. Peter Stafford, and Dr. Ned Field, thank you for timely providing requested data and for your expert opinions.

Dr. Misko Cubrinovski and Ms. Liz Ackerley are also thanked for their roles in providing me with the opportunity of lecturing undergraduate and postgraduate courses in the Civil and Mathematics departments, respectively. The computational resources provided by the Mathematics and Statistics department used for conducting the numerous analyses in this thesis were exceptional and I wish to thank Mr. Paul Brouwers and Mr. Steve Gourdie for allowing me this resource and teaching me how to use utilize its full potential.

Finally I would like to acknowledge the support of my family. To my sister, Lauren, and my partner, Brenda, thank you for all your love and support during the day-to-day events of life. To my mother, Gail, and father, Peter, thank you for providing me with all the

opportunities over the past 23 years which have enabled me to get to write these final words, this thesis is dedicated to you.

Brendon Bradley

17th April 2009

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Motivation

Consideration of seismic risk to structures can be viewed as five distinct steps: (i) spatial and temporal characterization of earthquake sources (faults); (ii) seismological prediction of earthquake-induced ground motions on outcropping rock and soil deposits; (iii) earthquake engineering prediction of the response of soil-structure systems; (iv) estimation of seismic response induced damage to components of the system; and (v) economic and societal implications of repairing damage, human injuries, and inoperability of the structure.

When considering these five aspects of seismic risk, the guiding concept is the age-old management phrase “what gets measured, gets managed”. That is, in order to make rational decisions regarding the seismic performance of a specific structure, quantifiable measures are needed for prioritization and allocation of resources.

The diverse range of expertise required to consider the aforementioned five steps in the earthquake problem necessitates that these steps be attained by various expert personnel. Thus, the analyst compiling data for use in determining seismic risk is left with the question of the relative reliability of the information and data obtained by each of the various expert personnel. The questionable reliability of data provided by seismic experts can be attributed to the complexity of the earthquake phenomena and its associated consequences. This complexity is reflected in uncertainties due to apparent phenomenological randomness and a lack-of-knowledge. Considering the magnitude of these uncertainties, it is prudent that they be explicitly considered in seismic risk computations.

The above paragraphs illustrate that three aspects are critical for a successful seismic risk assessment methodology. Firstly, the methodology should easily convolve the variety of expert knowledge and data encompassing the seismic risk problem. Secondly, the

methodology must provide quantitative measures of seismic performance which are insightful for the management of engineered structures. Finally, the methodology must account for the numerous and significant uncertainties which exist in the seismic risk assessment of structures.

In response to perceived insufficiencies of contemporary seismic performance and risk assessment methodologies the Pacific Earthquake Engineering Research (PEER) centre was founded in 1997 as a 10 year NSF-funded consortium of US universities and building-industry partners to "develop and disseminate technologies to support performance-based earthquake engineering" [1]. One of the two key developments of PEER was the development of the *PEER framework formula*. While in concept, the PEER framework formula can be applied at both structure-specific and region-specific levels, only the former is discussed in this dissertation.

The terminology in the PEER framework formula involves: decision variables (DV's), damage states (DS's), engineering demand parameters (EDP's) and intensity measures (IM's). DV's are quantitative measures which can be used to objectively assess seismic performance. DS's define the different discrete states of damage which a specific component in the structure is in following an earthquake, which in turn require specific repair tasks. EDP's are used to describe the seismic response of the structure, which may be in terms of peak or transient displacements and/or accelerations. IM's define the intensity (or severity) of the earthquake hazard to the structure. More often than not, IM's are used to define the severity of the earthquake-induced ground motion hazard. While the PEER framework formula is relatively new, several framework modifications have occurred since its initial presentation. Details of the initial developments are given in, for example, Aslani [2].

Using the four variables presented in the previous paragraph, a probabilistic formulation for computation of DV's based on the other three interim variables is given by:

$$\lambda_{\boldsymbol{DV}}(\boldsymbol{dv}) = \iint_{EDP,\,IM} \sum_{DS} G_{\boldsymbol{DV}|\boldsymbol{DS}}(\boldsymbol{dv}|\boldsymbol{ds}) P_{\boldsymbol{DS}|\boldsymbol{EDP}}(\boldsymbol{ds}|\boldsymbol{edp}) f_{\boldsymbol{EDP}|\boldsymbol{IM}}(\boldsymbol{edp}|\boldsymbol{im}) \left|\frac{d\lambda_{\boldsymbol{IM}}(\boldsymbol{im})}{d\boldsymbol{IM}}\right| d\boldsymbol{IM} d\boldsymbol{EDP} \quad (1\text{-}1)$$

where $\lambda_{\boldsymbol{X}}(\boldsymbol{x}) = \lambda_{\boldsymbol{X}}(\boldsymbol{X} > \boldsymbol{x})$ is the annual rate/frequency of $\boldsymbol{X}$ exceeding $\boldsymbol{x}$; $G_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X} > \boldsymbol{x}|\boldsymbol{Y} = \boldsymbol{y})$ is the complementary cumulative distribution (or survival) function (CCDF) giving the probability of $\boldsymbol{X}$ exceeding $\boldsymbol{x}$ given $\boldsymbol{Y} = \boldsymbol{y}$. $P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X} = \boldsymbol{x}|\boldsymbol{Y} = \boldsymbol{y})$ is the probability of $\boldsymbol{X} = \boldsymbol{x}$ given $\boldsymbol{Y} = \boldsymbol{y}$ (i.e. X is a discrete

variable in this case); and $f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{Y}=\boldsymbol{y}) = -\frac{dG_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y})}{d\boldsymbol{X}}$ is the probability density function of $\boldsymbol{X} = \boldsymbol{x}$ given $\boldsymbol{Y} = \boldsymbol{y}$. Note that a probability-based formulation equivalent to Equation (1-1) (which is frequency/rate-based) is employed in some chapters of this dissertation. In this case, the ground motion hazard is given by $G_{\boldsymbol{IM}}(\boldsymbol{im})$ such that $\left|\frac{dG_{\boldsymbol{IM}}(\boldsymbol{im})}{d\boldsymbol{IM}}\right| = f_{\boldsymbol{IM}}(\boldsymbol{im})$ and the result of the equivalent form of Equation (1-1) is $G_{\boldsymbol{DV}}(\boldsymbol{dv})$ which is the CCDF of the decision variable (for some specified time interval). Boldface font is used to illustrate that variables can potentially be vectors, although in many cases scalar variables are used for simplicity.

Equation (1-1) has several features worthy of note. Firstly, despite its apparent complexity, it is merely an application of the total probability theorem [e.g. 3]. Secondly, the assumption of conditional independence is made (e.g. $G_{\boldsymbol{DV}|\boldsymbol{DS}}(\boldsymbol{dv}|\boldsymbol{ds})$ implies that given $\boldsymbol{DS} = \boldsymbol{ds}$, the distribution of $\boldsymbol{DV}$ is independent of $\boldsymbol{EDP}$ and $\boldsymbol{IM}$). Thirdly, as a consequence of the conditional independence assumption, Equation (1-1) is comprised of the four independent tasks: (i) loss estimation ($\boldsymbol{DV}|\boldsymbol{DS}$); (ii) damage estimation ($\boldsymbol{DS}|\boldsymbol{EDP}$); (ii) demand estimation ($\boldsymbol{EDP}|\boldsymbol{IM}$); and (iv) hazard estimation ($\boldsymbol{IM}$).

It should become clear throughout this dissertation that Equation (1-1) and its variants provide a holistic framework for seismic performance and risk assessment. In particular, the PEER framework formula satisfies the three aforementioned critical aspects for a successful risk assessment methodology: (i) easy convolution of various expert knowledge and data; (ii) quantitative measures of seismic performance providing insight for the management of engineered structures; and (iii) explicit treatment of the uncertainties.

1.2 Objectives

Based on the aforementioned issues regarding seismic risk analysis and, in particular, the PEER framework formula this dissertation aims to address a variety of issues, namely:

1.2.1 Scrutinize simplified seismic hazard and risk methods

Several methods, predominantly employing closed-form solutions to variations of the PEER framework formula given by Equation (1-1) exist. Central to these closed-form solutions is the use of a power-law parametric form for the ground motion hazard in log-log

space. The effect of this assumption on the errors in the computation of demand and collapse hazards is addressed as well as an improved parametric form for the ground motion hazards which can be used to obtain semi-parametric solutions to the risk equations.

1.2.2 Technical advancement of the PEER framework formula and applications

While conceptually the PEER framework formula in itself provides a complete framework for PBEE, there are many developments which are still necessary in its application. From a technical standpoint this dissertation focuses on furthering several aspects of the framework, each discussed in the paragraphs below.

Selection of ground motion intensity measures for use in the PEER framework formula has to date revolved primarily around the prediction of peak displacement/drift demands in structures founded on stiff soils (such that soil response is not modeled explicitly). Attention in this dissertation is therefore given to the selection of intensity measures for use in seismic loss estimation, where loss is dependent on both acceleration and displacement response of the structure and its contents at spatially distributed locations; and structures supported on pile foundations embedded in liquefaction-susceptible soils, in which non-linear dynamic soil response and soil-structure-interaction are pivotal in the system deformational response.

General solutions of the PEER framework formula require the use of numerical methods for solving the integral equations. Despite the ever-increasing computational resources available, it is illustrated that solution of the PEER framework formula in the case of loss estimation of a real structure can be very computationally demanding and therefore it is prudent to scrutinize the numerical algorithms used in the seismic risk analysis. A numerical integration algorithm, termed magnitude-oriented adaptive quadrature (MAQ) is developed as an extension of general-purpose numerical quadrature algorithms, and is specifically tailored for efficient (i.e. a high ratio of accuracy vs. computational demand) solution of the integral equations in the PEER framework formula.

The significant computational demands of performing seismic risk analyses using the PEER framework formula also suggests the possible use of simplified methods of uncertainty propagation, which can reduce computational demands and be of sufficient accuracy in the case of preliminary seismic risk assessments. The efficacy of an uncertainty propagation method known as the first-order second moment (FOSM) method for use in the PEER framework formula is addressed. Particular attention is given to the trade-off between

accuracy and reduction in computational demand of the approximate solution relative to solution via direct numerical integration.

When a seismic risk assessment using the PEER framework formula comprises the seismic risk to multiple components, then shared uncertainties, or correlations, between the components have a significant effect on the results of the seismic risk assessment. To date the consideration of correlations has been limited by methodological tractability, numerical algorithms, and a paucity of data for their computation. A tractable and computationally efficient methodology for handling correlations between multiple components is developed and attention is given the determination of correlation coefficients, particularly in the case of insufficient empirical data.

In the PEER framework formula, given in Equation (1-1), the distributions of random variables which the integrations are performed over are those due to inherent randomness or aleatory uncertainty in each of the various relationships. Another significant source of uncertainty is that due to a lack-of-knowledge or epistemic uncertainty. While epistemic uncertainties are routinely considered in probabilistic seismic hazard analysis (PSHA), typically only the mean seismic hazard curve is used in seismic risk assessments. The consideration, characterization, and propagation of seismic hazard epistemic uncertainties in seismic risk analysis is discussed by rigorously examining epistemic uncertainties in the seismic hazard for various sites and intensity measures in the San Francisco bay area.

The development of component fragility functions is also burdened by significant epistemic uncertainties, which are generally not respected in fragility functions published in earthquake engineering literature. The consideration, quantification, combination, and propagation of various sources of epistemic uncertainties in fragility functions are addressed.

1.2.3 Application of the PEER framework formula for pile founded structures

In the first five years since the publication of the PEER framework formula research into its development and application have been, by and large, approached from the direction of structural engineering. The use of the PEER framework formula for soil-structure systems in which the effects of non-linear dynamic soil behaviour and soil-structure interaction are significant come with additional problems.

It is found that both for seismic risk analysis of structures with acceleration and displacement sensitive components, and seismic demand analysis of structures supported on

pile foundations embedded in liquefiable soils, that the ground motion intensity measure, spectrum intensity, *SI*, is a better predictor of seismic demand than other common IM's such as spectral acceleration at a single period. For routine use of such an intensity measure, high-quality ground motion prediction equations must be available for use in PSHA. A ground motion prediction equation for *SI* is developed based on ground motion prediction equations for spectral accelerations, which are available in most seismically-active regions of the world.

Intensity measures for the seismic response of soil-pile-structures, with particular focus on pile foundation demands is examined. A range of soil-pile-structure configurations are examined to determine those intensity measures which correlate well with the seismic demand on the pile foundation. The effects of bias due to ground motion selection and the influence of the kinematic and inertial mechanisms of pile deformation are also examined.

A state-of-the-art seismic risk assessment of a bridge structure is considered utilizing deaggregation-based ground motion selection, seismic effective stress analysis, and professional cost estimation, to obtain the seismic risk for repair cost and bridge inoperability decision variables. Deaggregation is used to illustrate the components of the bridge structure which are most vulnerable, thereby enabling clear emphasis on how the seismic performance of the system can be most effectively improved.

1.3 Organisation

This dissertation addresses several issues in seismic risk assessment. Chapter 2 illustrates the application of seismic risk assessment methods to a typical office building. Chapters 3 and 4 involve the scrutinization of simplified methods within the PEER framework formula. Chapters 5 through 11 are devoted to further development of various aspects of the seismic risk assessment methodology. Chapters 12 through 14 develop necessary prerequisites for the consideration of soil-pile-structure systems within the PEER framework formula

Chapter 2 focuses on the seismic risk assessment of a typical New Zealand office building. Attention is focused particularly on the vast variety of information which rigorous seismic loss estimation provides for effective decision making regarding seismic risk. The concept of loss deaggregation is illustrated as an effective method by which the contribution of various components in the structure toward the total seismic risk can be assessed. The use of seismic risk assessment results in decision making is also briefly illustrated.

Chapter 3 presents a parametric model for the seismic hazard curve, obtained from

probabilistic seismic hazard analysis. The parametric model, based on a hyperbola in $\log\lambda_{IM}$ – $\log IM$ space, is compared to the accuracy of the contemporary power-law model, which is linear in $\log\lambda_{IM}$ – $\log IM$ space. The fit of the hyperbolic hazard model to probabilistic seismic hazard analysis data for various intensity measures and locations in New Zealand is also illustrated. The primary use of the power-law model for the seismic hazard is that, when combined with a similar power-law form for the seismic demand-intensity relationship, it permits a closed-form solution for the annual frequency of exceeding a specified level of seismic demand (demand hazard). A semi-parametric closed-form solution for the demand hazard is developed using the hyperbolic hazard model and illustrated to be of superior accuracy than the power-law based solution.

Chapter 4 investigates the accuracy of the closed-form solution for the annual frequency of structural collapse (collapse hazard) which is based on a power-law form of the seismic hazard curve and a lognormal seismic capacity distribution. The effect of curvature (in $\log\lambda_{IM}$ – $\log IM$ space) of the seismic hazard curve, and the lognormal standard deviation of the collapse capacity are identified as two critical parameters influencing the error in the closed-form solution for the collapse hazard compared with direct solution via numerical integration. A parametric study on the magnitude of the error as a function of these two parameters for various methods of fitting the power-law seismic hazard curve, are investigated.

Chapter 5 investigates the efficacy of various ground motion intensity measures (IM's) in the prediction of spatially distributed seismic demands (EDP's) within a structure. To date, selection of intensity measures for seismic performance/risk assessment has been primarily focused on the prediction of peak interstorey drift in structures. This work has direct implications to building-specific seismic loss estimation, where the seismic demand on different components can be either sensitive to acceleration or displacement demands and is dependent on the location of the component in the structure. Several common intensity measures are investigated in terms of their ability to predict the spatially distributed demands in a 10-storey office building, which is measured in terms of maximum interstorey drift ratios and maximum floor accelerations. The efficiency, sufficiency and predictability of the various IM's are assessed for the various demand measures in the structure.

Chapter 6 further extends the work conducted in Chapter 5 by examining the effects of IM selection on performance assessment metrics such as: probability of structural collapse; probability of exceeding a specified level of demand or direct repair cost; and the distribution of direct repair loss for a given level of ground motion. A method is proposed to account for

the effect of varying seismological properties (e.g. magnitude, distance) of ground motions on seismic demand that does not require different ground motion records to be used for each intensity level.

Chapter 7 is devoted to the development of an efficient numerical integration algorithm for solving the integral equations of the PEER framework formula. The algorithm, called Magnitude-oriented Adaptive Quadrature (MAQ) is an integration algorithm with both locally and globally adaptive capabilities. It is illustrated how MAQ allows efficient integration over the entire integration domain and requires only an error tolerance and maximum number of function evaluations to be specified. The advantages of utilizing the MAQ algorithm over other conventional integration methods such as Romberg integration and conventional adaptive quadrature are illustrated for the numerical computation of (1) expected annual loss; and (2) annual rate of collapse.

Chapter 8 focuses on the efficacy of an approximate method of uncertainty propagation, known as the first-order second-moment (FOSM) method, for use in seismic loss estimation. The governing probabilistic equations which define the PEER-based loss estimation methodology used are discussed, and the proposed locations to use the FOSM approximations identified. The justification for the use of these approximations is based on a significant reduction in computational time by not requiring direct numerical integration, and the fact that only the first two moments of the distribution are known. Via various examples it is shown that great care should be taken in the use of such approximations, particularly considering the large uncertainties that must be propagated in a seismic loss assessment. A complete loss assessment of a structure is considered to investigate in detail the location where significant approximation errors are incurred, where caution must be taken in the interpretation of the results, and the computational demand of the various alternatives.

Chapter 9 addresses statistical dependencies, or correlations, between multiple components in structure-specific seismic loss estimation. To date, the consideration of such correlations has been limited by methodological tractability; increased computational demand; and a paucity of data for their computation. The effect of component correlations, which arise in various forms, is however a significant factor affecting the results of structure-specific seismic loss estimation and therefore it is prudent that adequate consideration is given to their effect. Details of a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered are developed and discussed. Methods to determine the necessary correlations are discussed, particularly those that can be used in the absence of sufficient empirical data, for which values are suggested based on judgement. The

effects of various assumptions regarding correlations are illustrated via application to a case-study office structure.

Chapter 10 investigates epistemic uncertainties in seismic hazard analyses and their role in the broader picture of seismic performance assessment of structures and facilities. Using the 2002 Working Group on California Earthquake Probabilities earthquake rupture forecast for the San Francisco bay area, the epistemic uncertainty for several different intensity measures at several sites is investigated. Normalization of the epistemic uncertainty in the seismic hazard for various sites and ground motion intensity measures is used to investigate trends in the uncertainty magnitude as a function of probability of exceedance. The distribution and correlation structure of the epistemic uncertainty is investigated as well as various methods of propagation using an example of the collapse hazard of a structure.

Chapter 11 is concerned with the inclusion of epistemic uncertainties in component fragility functions used in performance-based earthquake engineering. Conventionally fragility functions, defining the probability of incurring at least a specified level of damage for a given level of seismic demand, are defined by a mean and standard deviation and assumed to have a lognormal distribution. However, there exist many uncertainties in the development of such fragility functions. The sources of epistemic uncertainty in fragility functions, their consideration, combination and propagation are discussed and illustrated using examples presented in literature. The benefits of the consideration of epistemic uncertainties pertaining to the documentation, quality assurance, implementation, and updating of fragility functions are discussed.

Chapter 12 focuses on the development of a ground motion prediction equation for spectrum intensity, *SI*. *SI*, defined as the integral of the pseudo spectral velocity of a ground motion from 0.1 to 2.5 seconds, is shown in Chapters 5, 6, and 13, to be an intensity measure that efficiently predicts the seismic response of structures for seismic risk assessment as well as the seismic response of liquefiable and non-liquefiable soil deposits and the seismic demands on pile foundations. As such relationships developed specifically for *SI* are sparse, the development of a relationship based on current ground motion prediction relations for spectral acceleration, which are available in most regions of seismic activity is illustrated. Rigorous statistical tests are performed to confirm approximations regarding the distribution of the prediction equation. The lognormal standard deviation of the *SI* prediction equation is compared and contrasted to other common ground motion intensity measures.

Chapter 13 investigated the efficacy of various ground motion intensity measures for the seismic response of pile foundations embedded in liquefiable and non-liquefiable soils. A

soil-pile-structure model consisting of a two layer soil deposit with a single pile and a single-degree of freedom superstructure is used in a parametric study to determine the salient features of the seismic response of the soil-pile-structure system. A suite of ground motion records scaled to various levels of intensity are used to investigate the full range of pile behaviour, from elastic response to failure. Various intensity measures are used to inspect their efficiency in predicting the seismic demand on the pile foundation for a given level of ground motion intensity. The effects of kinematic and inertial mechanisms in the soil-pile-structure response and their effect of the efficiency of intensity measures are also discussed.

Chapter 14 presents the seismic performance assessment of a New Zealand bridge structure supported on pile foundation embedded in liquefiable soils. Seismic effective stress analysis using advanced constitutive models is used to rigorously examine the seismic response of the system including complex non-linear dynamic soil response and soil-structure interaction. Deaggregation of the seismic hazard is used to select ground motion records for use in assessing the seismic response under multiple ground motions at multiple levels of seismic intensity. Based on the mechanical properties of the components and professional cost estimation, fragility and loss functions for each of the components are obtained and the seismic risk assessment methodology developed throughout this dissertation is employed to rigorously assess the seismic performance and risk to the bridge structure, in terms of both direct repair cost and loss of functionality consequences.

Chapter 15 discusses the key contributions of this dissertation in the field of seismic risk assessment. Limitations in the contributions presented which can be relaxed and extended in future are also recommended.

This dissertation has been written as a collection of chapters which are intended to be stand-alone publications. This has two repercussions for readers. Firstly, some introductory material in each of the chapters, although examined in specific reference to the logic in the chapter, is repeated from earlier chapters. Secondly, the notation used, while consistent within a single chapter, is not necessarily consistent with the remainder of the dissertation. Apologies are made to those intending to read this dissertation as a continuous document.

1.4 References

[1] PEER. Pacific Earthquake Engineering Research centre: Mission and goals. http://peer.berkeley.edu/about/mission_goals.html. Last accessed: 25 February 2009

[2] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[3] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

2. Seismic Loss Estimation for Efficient Decision Making

Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society of Earthquake Engineering* 2009; **42**(2): 96-110.

2.1 Abstract

In order to incorporate seismic risk of facilities into a decision making framework, procedures are needed to quantify such risk for stakeholders. Seismic loss estimation methods combine seismic hazard, structural response, damage fragility, and damage consequences to allow quantification of seismic risk. This chapter presents a loss estimation methodology which allows various means of quantifying seismic risk of a specific facility. The methodology is component-based and can therefore distinguish between different structural configurations or different facility contents and is consistent with state-of-the-art loss assessment procedures. Loss is measured in the forms of direct structural and non-structural repair costs, and although not considered in the example, business disruption and occupant injuries can also be considered. This framework has been packaged in a computer code available for future dissemination in the public domain so that users need only to have a basic understanding of the methodology and the input data that is required. Discussion is given to the flexibility of the framework in terms of the rigour which can be employed at each of the main steps in the procedure. Via a case study of a high-rise office building, the use of the methodology in decision-making is illustrated. Methodological requirements and further research directions are discussed.

2.2 Introduction

Current seismic design codes provide guidelines for the design and detailing of structures with the primary goal of preventing global collapse during strong ground motion shaking. Observations from worldwide earthquakes in the past two decades have illustrated that with few exceptions, structures designed to these state-of-practice guidelines are sufficient for providing life safety with a high degree of confidence. However, these events illustrated the severe economic consequences resulting from earthquakes in highly developed regions of society. These economic consequences can be primarily attributed to: (i) direct economic losses associated with repairing damage within a structure; (ii) direct losses associated with injuries and casualties; and (iii) indirect losses associated with the loss of income due to business disruption. These three forms of losses (damage, death and downtime) are known as the '3D's'. Some examples from the United States include the 1994 Northridge ($17-26 billion), and 1989 Loma Prieta ($11 billion) earthquakes [1]. In response to these observed losses it has become apparent that seismic design of structures should consider all of these potential consequences and their likelihood of occurrence.

Quantification of seismic risk is a difficult task which is subject to inherent variability. Although it can be roughly forecasted, it is not known when and where future significant fault ruptures will occur. Even when an earthquake occurs at a particular location, due to the complex rupture mechanism, process of seismic wave propagation and site effects, the intensity, frequency content and duration of ground motions at a particular site are also uncertain. These uncertainties in the ground motions affecting a given site cause corresponding uncertainty in the level of structural response, and associated damage in the structure. Finally, the cost to repair a damaged structure is also uncertain and depends on available resources and demand. In order to rigorously assess the seismic risk of a structure all of the above uncertainties should be accounted for. Thus it becomes necessary that the problem of seismic risk is cast into a probabilistic framework which can propagate such uncertainties in each of the input variables and give a probabilistic output useful for decision making processes.

These aforementioned uncertainties result from either inherent randomness in a process, or uncertainty due to the limited knowledge and application of engineering models. These two different types of uncertainty are referred to as aleatory randomness and epistemic uncertainty, respectively. An example of aleatory randomness would be the variation in the

level of ground motion observed at a site due to different ground motions resulting from the same rupture magnitude and source-to-site distance, while epistemic uncertainty would result from the ground motion prediction equation used to estimate the level of ground motion at the site. As aleatory randomness is deemed as an inherent property of complex phenomena, it cannot be reduced; epistemic uncertainty being knowledge-based can be reduced if better knowledge of the phenomena is acquired. As these two different uncertainties are related to different aspects of the considered problem they deserve separate treatment within a decision making process.

This chapter presents a discussion of the use of seismic loss estimation for decision making at various stages of design and/or seismic assessment of structures. An outline of the state-of-the-art seismic loss estimation is given with reference to a specific case study of a 10-storey New Zealand commercial office building. Using the case study structure a full loss assessment is performed and discussion is given to each of the possible outputs for decision making. Some simplistic assumptions are made in the loss assessment in order to accomplish the goal of this chapter which is to present the interpretation of loss assessment results for use in decision making. For simplicity, this chapter is primarily concerned with aleatory randomness only. These assumptions are revisited at the end of the chapter where further discussion is given to the effects of relaxing such assumptions.

2.3 Case study structure

The case study structure used herein to illustrate the use of seismic loss estimation tools in decision making process is based on the Red Book building [2] which acts as a design example of the New Zealand Concrete Code [3]. Figure 2-1 illustrates plan and elevation views of the building layout. The primary lateral load carrying system consists of four one-way perimeter moment resisting frames which are 3 bays long. Vertical loads are transferred primarily through interior columns with gravity beams supporting one-way floor units. Although originally designed for a site in Christchurch, in this study it was assumed that the structure is located at a site in Wellington. The soil is assumed to be class A [4] and stiff enough so that local site effects are not significant in modifying the bedrock ground motion.

A 2D model of perimeter frame was developed using the time-history analysis program Ruaumoko2D [5]. Due to the symmetry of the structure, it was assumed that the 3D response could be reasonably approximated by separate 2D analyses in each of the two primary directions. A fixed-base model was used in the analysis and as a result soil-structure-

foundation interaction was neglected. The structure was modelled using a lumped mass model and non-linear (beam) elements based on the modified Takeda hysteresis, with the appropriate section properties determined using fibre-based biaxial section modelling. The structural model had a fundamental period of 1.5 seconds. Further details on the structural response and the effects of the assumptions on the loss estimation outcome are discussed later in the chapter.

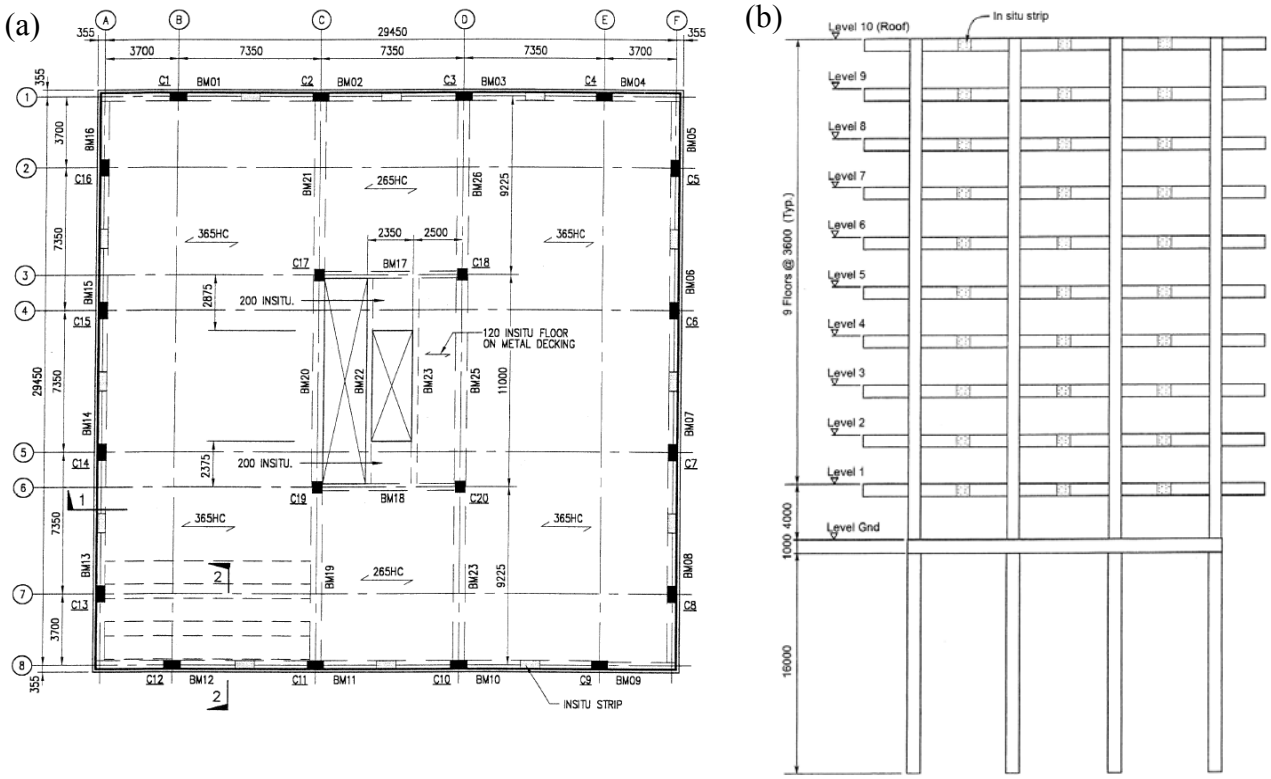


Figure 2-1: (a) Plan; and (b) elevation of the Red Book building [2].

2.4 Seismic hazard

The seismic hazard at the site of the structure can be quantified by performing a probabilistic seismic hazard analysis (PSHA). PSHA combines the magnitude recurrence relationships of various earthquake sources, and a ground motion prediction relationship. The ground motion prediction relationship describes the level of ground motion shaking at a site as a function of the magnitude of the earthquake and faulting type; source-to-site distance and path effects; local site effects; and soil amplification. The result of a PSHA is a ground motion hazard curve which gives the annual frequency of exceeding specific values of ground motion intensity. In this study, the ground motion hazard for Wellington based on Stirling *et al*. [6] was employed. As the fundamental period of the structure is 1.5 seconds, the 5% damped spectral acceleration at this period, S_a(T=1.5s,5%) (or simply S_a for brevity) is chosen

as the ground motion intensity measure (IM). This selection of ground motion IM is based on the observation from past researchers [e.g. 7] that the spectral acceleration at the fundamental period of the structure is an 'efficient' IM at predicting the drift demands in the structure. An efficient IM is desired as a reduction in the uncertainty in the structural response will result in a reduction in the uncertainty in the magnitude of economic losses. Figure 2-2 illustrates the seismic hazard curve obtained via PSHA which (in this case) describes the mean annual frequencies of exceeding various levels of S_a. The two reference lines indicated in the plot represent 10% and 2% probability of exceedance in 50 years (return periods of 475 and 2475 years, respectively) which are typical frequencies used when determining forces/displacements for the design and assessment of structures.

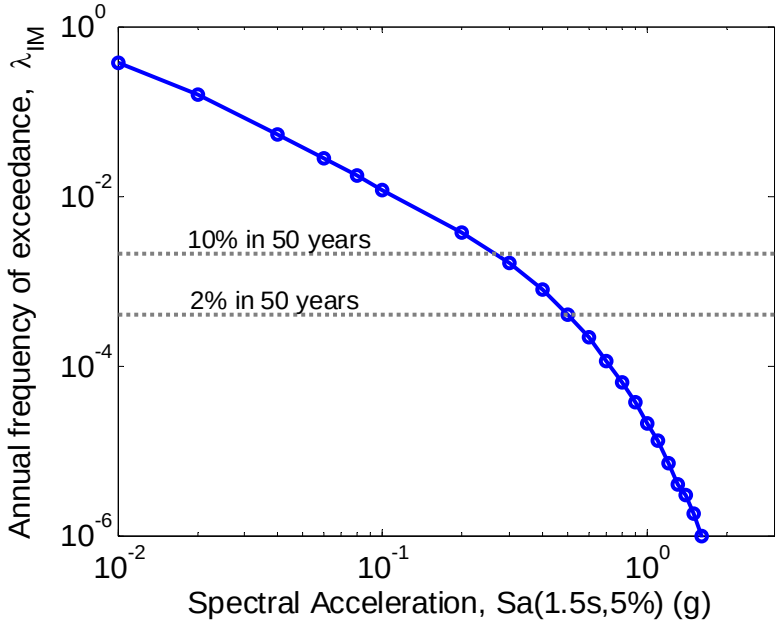


Figure 2-2: Ground motion hazard curve for Wellington, New Zealand [6].

2.5 Seismic response analysis

Seismic response analysis of the structure was performed using non-linear time history analysis with a suite of ground motions scaled over a wide range of ground motion intensities to account for the variability in structural response due to differences in ground motions of the same intensity (termed record-to-record randomness). A suite of 40 ground motion records complied by Medina and Krawinkler [8] were used for conducting the non-linear time history analyses. The suite contains ground motions recorded on stiff soil with magnitude and distance ranges of 6.5-6.9 and 13.3-39.3 km, respectively. The suite is termed 'ordinary' by Medina and Krawinkler, as none of the records show effects of near-fault motions (i.e.

directivity or 'fling' effects), and all ground motions were recorded on stiff soils. As the Wellington seismic hazard will likely have some dominance due to near-fault records, in a more rigorous assessment the suite of ground motions should incorporate near-fault ground motions (which more accurately match the likely spectral shapes of future ground motions), however this is beyond the scope of this chapter.

In the example presented herein, the seismic response is evaluated rigorously using time history analyses with multiple ground motions. In preliminary investigations however, it may not be feasible to conduct such rigorous analyses. In such cases there are various simplified methods which may be used to approximate the 'true' seismic response obtained by direct non-linear time-history analysis. Such methods can be of a pseudo-static nature or modal pushover methods coupled with single-degree-of-freedom (SDOF) time history analyses (e.g. [9-12]).

The suite of ground motions were scaled to S_a values ranging from 0.1-1.5g in increments of 0.1g. Thus in total, 600 non-linear time history analyses were completed using 40 different ground motion records at 15 different intensities. This process of incrementally carrying out time history analysis using ground motion records to various levels of intensity is termed Incremental Dynamic Analysis (IDA) [13]. IDA allows quantification of the seismic demand and capacity of the structure from initial elastic response through to global collapse. Herein the interstorey drift (or drift for brevity) between two floors is denoted as the interstorey drift occurring in the lower of the two floors (i.e. the second floor interstorey drift is the interstorey drift between the second and third floors). This terminology is adopted to allow clarity in later discussions considering loss due to both drift and acceleration demands. Figure 2-3 illustrates two IDA results representing the maximum interstorey drifts (the Engineering Demand Parameter, EDP) on the second and eighth floors of the analysed structure as a function of the ground motion IM. Each of the points in the figures is the result of a single time history analysis when global structural collapse did not occur, while the two lines indicate the mean response for a given level of intensity with and without consideration of global collapse, respectively (discussion on computing these two quantities is given further in the chapter). The fewer number of analysis points in the figure at higher levels of ground motion intensity is due to the fact that a larger proportion of ground motions cause collapse (and are therefore not displayed).

As is evident in Figures 2-3a and 2-3b there exists a significant amount of variation in the seismic response of the structure for various levels of ground motion, with the variation increasing as the level of intensity increases (primarily due to significant non-linear response

of the structure). This uncertainty can be quantified by computing the dispersion (standard deviation of the logarithm of the analysis data where collapse does not occur) of the responses as a function of intensity which is shown in Figures 2-3c and 2-3d. Dispersion is used as the measure of variation as it has been shown by various researchers that this variation can be

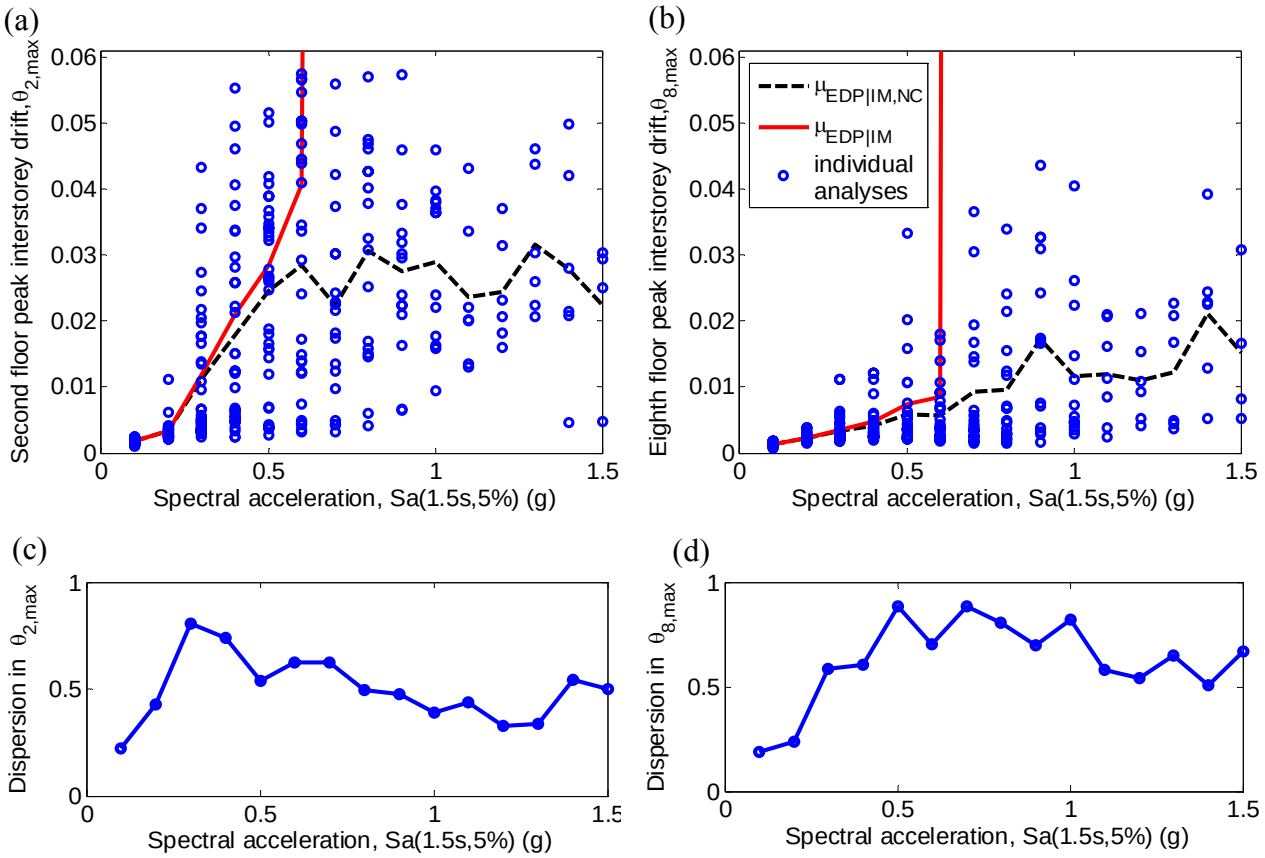


Figure 2-3: Examples of IDA plots of maximum interstorey drift for: (a) second-third storey; (b) seventh-eighth story; (c) dispersion in second-third storey drift response; and (d) dispersion in seventh-eighth storey drift response.

well approximated by the lognormal distribution [e.g 14, 15]. The dispersion, β, is related to the non-log moments of the variable by [16]:

$$\beta^2 = \sigma_{\ln X}^2 = \ln\left(\left[\frac{\sigma_X}{\mu_X}\right]^2 + 1\right) \tag{2-1}$$

where μ_X and σ_X are the mean and standard deviation of the variable X. For reference, it can be shown that a first order (Taylor series) approximation of Equation (2-1) is $\beta \approx \delta$, where $\delta = \sigma/\mu$ is the coefficient of variation (cov).

With the mean and dispersion of the seismic response (given no collapse) known, for a given level of intensity, the probability of exceeding some level of demand can be obtained

from:

$$G_{EDP|IM}(edp|im) = G_{EDP|IM,NC}(edp|im)\left[1 - P_{C|IM}(im)\right] + P_{C|IM}(im) \quad (2\text{-}2)$$

where $G_{X|Y}(x \mid y)$ is given as shorthand notation for the complementary cumulative distribution function (CCDF) of x given y, $G(X > x \mid Y = y)$ [16]; $G_{EDP|IM,NC}(edp|im)$ gives the probability of $EDP > edp$ given $IM = im$ and no collapse, and $P_{C|IM}(im)$ is the probability of collapse given $IM = im$ (defined in the subsequent section). The mean demand considering both collapse and non-collapse cases (shown in Figures 2-3a and 2-3b) can be computed indirectly from Equation (2-2). Specific details on this computation can be found in Jalayer [17].

The variation in the seismic demand at different elevations is illustrated in Figure 2-4 which shows the (mean) peak interstorey drifts and (mean) peak floor accelerations as a function of the elevation in the building. Note that the values in the figure represent the mean of many ground motion records and each of the different values will not likely occur at the same time (i.e. these are not profiles at a specific step in time). As is typical for a multi-storey frame structure the drift demands are observed to be the most severe in the lower half of the structure, while the total (as opposed to relative) acceleration demands are approximately constant over the height of the structure for low levels of shaking, but become larger in the lower stories for higher levels of ground motion when significant damage in the lower floors occurs (i.e. the damaged floors begin to isolate the upper region of the structure). It will be shown later in the chapter that this localization of drift and acceleration demands at high levels of ground motion in the lower stories of the structure prevents significant damage to components in the upper stories.

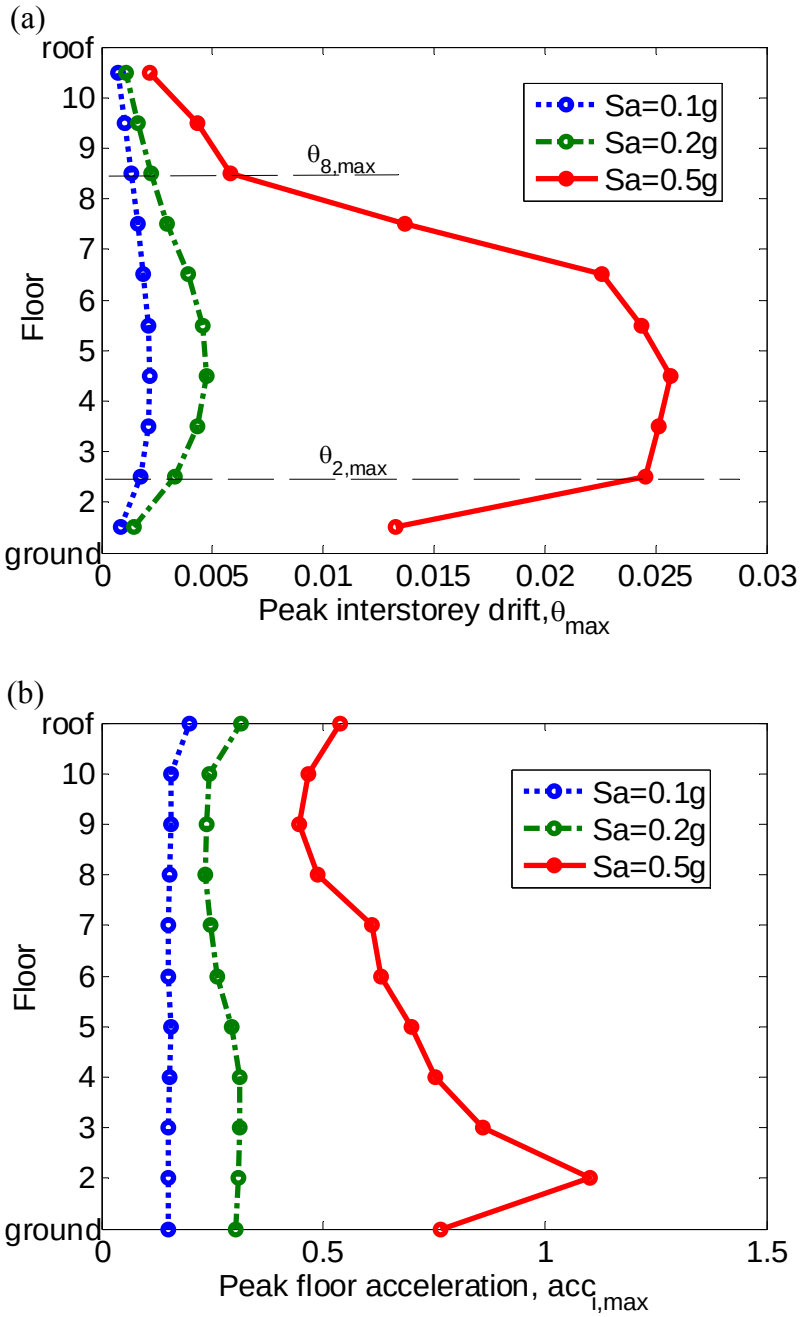


Figure 2-4: Variation in: (a) mean drift; and (b) mean acceleration demands over the height of the structure.

2.6 Occurrence of global collapse

In loss assessments as well as for decision making purposes it is beneficial to consider separately global collapse cases and non-collapse cases [18]. From a loss assessment perspective, this is done as the loss due to global collapse can not simply be obtained by summing the losses for each component within the structure (which is the case when global collapse does not occur). This is because post-earthquake reconstruction for a structure which

has sustained global collapse generally involves demolition of the structure and replacement by a new one, and thus some components which may not be damaged are nevertheless replaced. From a seismic response viewpoint, if global collapse is not considered separately then the assumption of the *EDP*|*IM* relationship having a lognormal distribution is likely to be violated [19]. Finally, by separately considering global collapse it is possible to determine the likelihood of global collapse for a given earthquake scenario or as a probability over the lifetime of the structure, which is an important measure in assessing seismic performance [20].

Traditionally, the occurrence of structural collapse has been associated with some prescribed level of seismic demand, such as interstorey drift or component plastic deformation. This however does not account for the redundancy of structural systems which allows for redistribution of damage and global stability despite local failures [20]. Here collapse is defined as the state in which sidesway instability occurs in one or more storeys. Collapse due to loss of vertical carrying capacity (LVCC) (due to axial and critical shear failures) is not considered due to a lack of structural analysis tools which can reliably capture these phenomena, although work is progressing rapidly in this area [21]. From IDA results a collapse fragility curve can be constructed by first determining the probability of collapse for various levels of ground motion intensity (based on the proportion of records which cause structural collapse), and then typically fitting these raw data points with a lognormal distribution. Figure 2-5 illustrates the collapse fragility curve for the case study structure. Note that analyses were only conducted up to 1.5g S_a, and thus extrapolation is needed to describe the probability of collapse at higher intensity levels. As will be shown later, the infrequent occurrence of such events (larger than 1.5g S_a) means that these levels of intensity do not significantly affect the results in the loss analyses. It is also noted that the hysteresis used for the structural elements in the Ruaumoko2D structural model considers strength and stiffness degradation but not cyclic degradation which can be significant for responses to ground motions of significant duration. It is also possible to account for collapse due to LVCC based on a weakest-link concept using the collapse fragilities of all structural components within the structure [19], but is not done so in this example. The two aforementioned points result in a collapse fragility curve shown in Figure 2-5 to be un-conservatively biased, however it is used here for the purposes of illustration.

In conceptual design phases it is not feasible to obtain the collapse fragility curves from time-history analysis. In such cases, it is recommended to use the results of Medina and Krawinkler, [8] and Ibarra and Krawinkler, [22] who have conducted extensive analysis on

the collapse capacity of frame and wall systems of various geometry and hysteretic behaviour, and provide simple estimates of the median level of intensity causing global collapse.

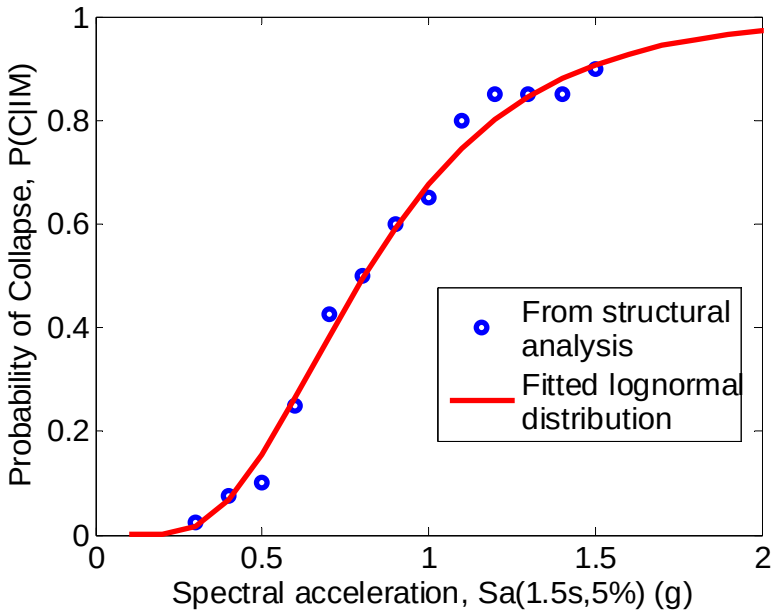


Figure 2-5: Collapse fragility curve for case study structure

2.7 Structural response and collapse hazards

The results of the seismic response analysis can be combined with the ground motion hazard to provide the rates of exceedance for various levels of seismic demand on the structure. For example, by combing the collapse fragility curve and the ground motion hazard, it is possible to compute the annual rate of global collapse by:

$$\lambda_C = \int P_{C|IM}(im) \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM \tag{2-3}$$

where λ_C is the annual rate of structural collapse; $P_{C|IM}(im)$ is the probability of collapse for a given $IM = im$ obtained from the collapse fragility curve (Figure 2-5) and $\lambda_{IM}(im)$ is the ground motion hazard curve (Figure 2-2). Equation (2-3) is simply an application of the total probability theorem and indicates that the annual rate of structural collapse can be obtained by multiplying the probability of collapse for a given *IM* with the likelihood of that level of *IM* occurring, and then summing over all the possible values of *IM*. As historical and recent earthquake reconnaissance indicates that structural collapse is the primary source of casualties and loss of life caused by earthquakes, then the annual rate of structural collapse is a key performance criterion for use in seismic assessment of structures. Also, since the primary purpose of current standards is to implicitly design structures that will provide life safety by

maintaining structural integrity, then Equation (2-3) provides an explicit method for assessment of this implicit objective of code-designed structures. For the case study structure considered in this chapter it was found that the annual rate of collapse is 2.15×10^{-4}, which (based on the Poisson assumption) corresponds to a 1.1% probability of global collapse of the structure over a service life of 50 years. As previously mentioned, the collapse fragility curve was unconservatively biased since LVCC and cyclic deterioration were ignored, and accounting for these two effects is likely to slightly increase the annual rate of collapse.

Similar to the annual rate of collapse, the seismic response for a particular EDP can be combined with the ground motion hazard curve to obtain the annual rate of exceeding various levels of *EDP* (herein referred to as an EDP hazard curve) by:

$$\lambda_{EDP}(edp) = \int G_{EDP|IM}(edp|im) \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM \tag{2-4}$$

where $\lambda_{EDP}(edp)$ is the annual rate of exceeding $EDP = edp$.

Figures 2-6a and 2-6b illustrate the EDP hazard curves for interstorey drifts and accelerations, respectively for the investigated structure. Several things should be noted from Figure 2-6 in relation to the results from the structural analysis (Figure 2-3-Figure 2-5) and the ground motion hazard (Figure 2-2). Firstly, all of the hazard curves have an asymptote at $\lambda = 2.15 \times 10^{-4}$ which is the annual rate of global collapse. Secondly, for a given level of exceedance rate the values for each of the lines (representing different floors) can be used to create figures similar to that of Figure 2-4, for different levels of annual rate of exceedance. Finally, the location of the lines relative to each other therefore describes how the distribution of seismic demand changes over the height of the structure as a function of ground motion intensity. For example, since all of the drift EDP hazard curves remain relatively parallel to each other it indicates that the hierarchy of interstorey drifts over the height of the structure remains relatively unchanged as the level of *IM* increases. However, for the acceleration EDP hazard curves it is immediately apparent that the curve for the roof acceleration crosses over the ground, 2nd and 6th floor acceleration curves as the rate of exceedance decreases. This

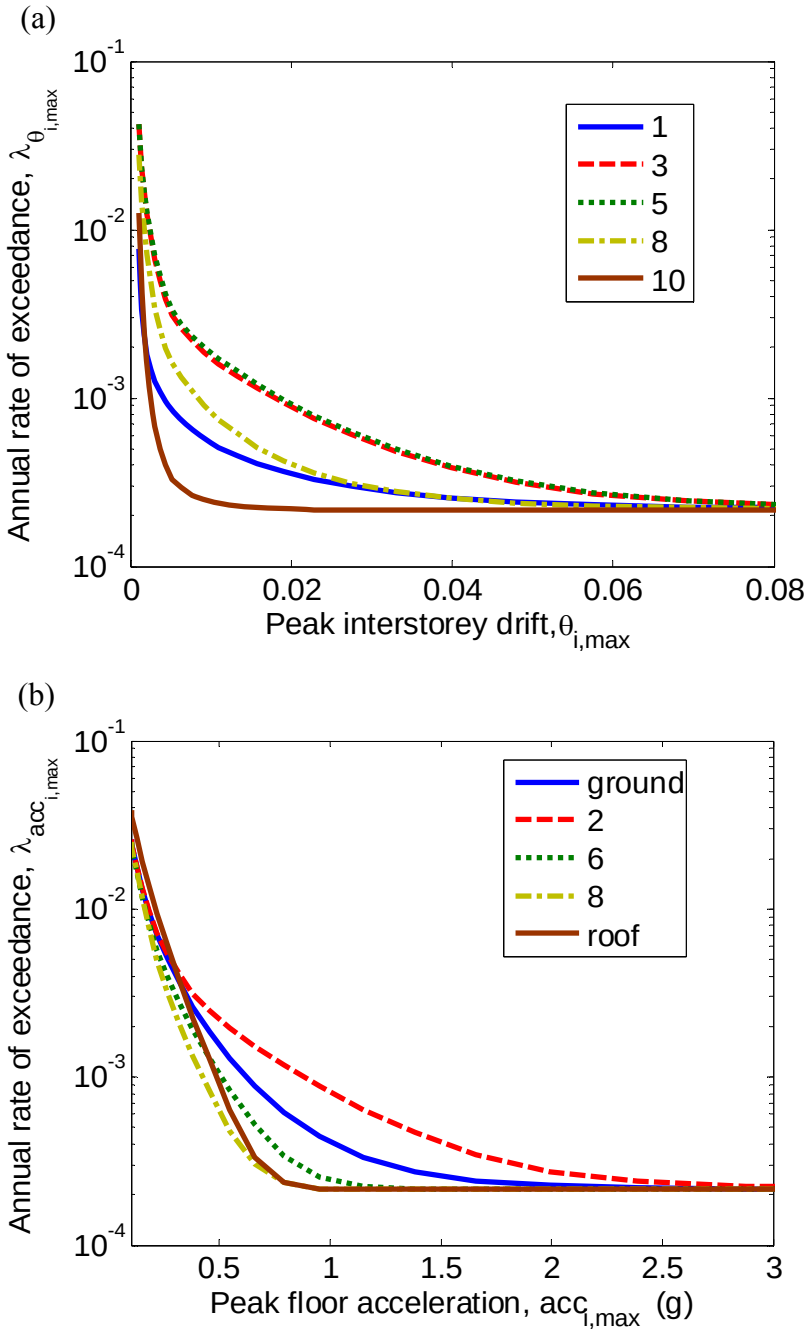


Figure 2-6: EDP hazard curves of: (a) peak interstorey drift; and (b) peak floor accelerations for the case study structure.

indicates that the roof acceleration which is larger than the maximum acceleration of other floors for small levels of ground shaking gradually becomes less than the maximum acceleration of several floors as the ground motion level increases (this same observation can be seen in Figure 2-4b as S_a increases).

2.8 Component inventory of structure

In order to carry out loss estimation for a specific structure, an inventory of components and their location within the structure is required. The level of detail regarding quantity estimates for each of the components within the structure will depend on the state which the structure is in when the assessment is carried out. In a retrofit assessment where all of the component quantities are known exactly is significantly different compared to the preliminary stage of assessing various structural design configurations. Table 2-1 gives a list of approximate densities of various non-structural components and contents in office buildings based on Aslani [19] ; Mitrani-Reiser [23]; and Buchan [24] which is likely to be useful in the latter design scenario.

Table 2-1: Typical densities of non-structural components and contents in office buildings.

Component	Description	Density
Interior partitions, finish, and paint	16mm 1-side partition on metal stud (same for finish).	0.8 m^2 Partition / m^2 floor area (FA) (Mitrani-Reiser, 2007)
Exterior glazing	1.5m x 1.8m standard glass panes	1.3 panes / m floor perimeter (Mitrani-Reiser, 2007)
Acoustical ceiling	0.6m x 1.2m tiles with Aluminium frames	0.8 tiles / m^2 FA (Mitrani-Reiser, 2007)
Automatic sprinklers	3.7m sections of sprinkler piping	0.08 sections / m^2 FA (Mitrani-Reiser, 2007)
Servers and network equip	Typical	\$235 / m^2 FA (Buchan, 2007)
Computers and printers	Typical	\$102 / m^2 FA (Buchan, 2007)
Bookcases and file cabinets	Typical	\$9 / m^2 FA (Buchan, 2007)
Roof mounted equipment	Coolers, airconditioning etc.	\$ 64 / m^2 FA (Buchan, 2007)
Workstation desks	Typical	\$24 / m^2 FA (Buchan, 2007)
Generic acceleration sensitive	fire protection systems, HVAC, Heating, cooling, pumps, plumbing, toilets	\$110 / m^2 FA (Aslani, 2005)
Generic drift sensitive	vertical piping, bath tubs, F.H.C, Ducts	\$110 / m^2 FA (Aslani, 2005)

When conducting a loss assessment of a structure, it is important to consider all of the components which have the potential to significantly contribute to the loss due to earthquakes causing a wide range of shaking intensities. Taghavi and Miranda [25] researched typical cost distributions for office, hotel and hospital buildings. In all three building types they found that each of the different component types (structural, non-structural and contents) are

important. For the office building considered in this study the components listed in the first column of Table 2-2 have been considered.

Table 2-1 was used to determine the quantities for the non-structural and contents components, while quantities for the structural components were obtained directly from building drawings (Figure 2-1). It was assumed that all of the contents and non-structural components had equal densities over the height of the building with the exception of the roof mounted equipment (located on the roof only), server and network equipment (located on the 3rd, 6th, and 10th floors), and elevators (ground floor). All information regarding the specifications of each of the various components can be found in Bradley [26].

Table 2-2: Quantities used in the case study example.

Component	Description	Quantity
Ductile eam-column joints	Post 1960s ductile beam column joints (2 beams)	24 / floor
Columns	Gravity columns (and seismic columns on first floor)	20 on 1st floor, 4 on all other floors
Slab-beam-column connections	Connection of slab to seismic frame	24 / floor
Partition	Drywall partitions and finish	721 m^2 / floor
Exterior glazing	1.5m x 1.8m standard glass panes	99 panes / floor
Acoustical ceiling	0.6m x 1.2m tiles with Aluminium frames	693 tiles / floor
Automatic sprinklers	3.7m sections of sprinkler piping	23 sections / floor
Servers and network equip	Typical	$260,000 on floors 3,6, and 10
Computers and printers	Typical	$93000 / floor
Bookcases and file cabinets	Typical	$16200 / floor
Roof mounted equipment	Coolers, airconditioning etc.	$600,000 on roof
Workstation desks	Typical	$21600 / floor
Generic acceleration sensitive	fire protection systems, HVAC, Heating, cooling, pumps, plumbing, toilets	$100,000 / floor
Generic drift sensitive	vertical piping, bath tubs, F.H.C, Ducts	$100,000 / floor

Each different component has its own fragility and loss functions which are defined based on the various damage states requiring repair that the component could be in following the earthquake. In this example loss was considered as that due to direct repair costs of replacing the structure to its original state and losses due to deaths and downtime were not

considered. Fragility and loss function for all of the components in Table 2-2 can be found in Bradley [26], and therefore only a single example is outlined here to illustrate the methodology.

2.9 Loss demand relationship

Figures 2-7a and 2-7b illustrate fragility and loss functions for the four damage states used to define the seismic performance of ductile reinforced concrete beam-column joints. For each of the damage states there is uncertainty (represented here by cumulative lognormal distributions) due to randomness in the strengths of the joints and the cost to repair the various levels of damage. Such fragility functions for structural components are typically obtained from a database of experimental results, although analytical methods and expert opinion can also be used [27], while the loss functions are typically obtained by employing a professional cost estimator [23].

The fragility and loss functions can be combined to compute the mean and variance of the loss for a given level of *EDP* by the following equations:

$$\mu_{L_i|EDP}(edp) = \sum_{j=1}^{N_{DS}} \mu_{L_i|DS_j}(ds_j) P_{DS_j|EDP}(ds_j \mid edp) \tag{2-5}$$

$$\sigma^2_{L_i|edp} = \sum_{j=1}^{N_{DS}} \left[\mu^2_{L_i|DS_j}(ds_j) + \sigma^2_{L_i|DS_j}(ds_j)\right] P_{DS_j|EDP}(ds_j \mid edp) - \mu^2_{L_i|EDP}(edp) \tag{2-6}$$

where $\mu_{L_i|EDP}(edp)$ and $\sigma^2_{L_i|EDP}(\,)$ are the mean and variance of the loss for given $EDP = edp$; $\mu_{L_i|DS_j}(ds_j)$ and $\sigma^2_{L_i|DS_j}(ds_j)$ are the mean and variance in the loss-damage state ($L|DS$) relationship (Figure 2-7b) ; N_{DS} is the number of damage states; and $P_{DS_j|EDP}(ds_j \mid edp)$ is the probability of being in damage state $DS = ds_j$ given demand level $EDP = edp$ which can be computed as the vertical distance between the fragility curves shown in Figure 2-7a:

$$P_{DS_j|EDP}(ds_j \mid edp) = \begin{cases} G_{DS_j|EDP}(ds_j \mid edp) - G_{DS_j|EDP}(ds_{j+1} \mid edp) & j \neq N_{DS} \\ G_{DS_j|EDP}(ds_j \mid edp) & j = N_{DS} \end{cases} \tag{2-7}$$

where $G_{DS_j|EDP}(ds_j \mid edp)$ is the probability of exceeding ds_j given demand level $EDP = edp$ (i.e. from the fragility curves in Figure 2-7a). Figure 2-7c illustrates the mean and mean ± one standard deviation values for the $L|EDP$ relationship for the RC joint, while Figure 2-7d

illustrates the dispersion in the loss as a function of *EDP*. The large dispersion for small *EDP* values apparent in Figure 2-7d is typical of *L*|*EDP* curves for various types of components. This is due to a large standard deviation in the loss and a small mean loss.

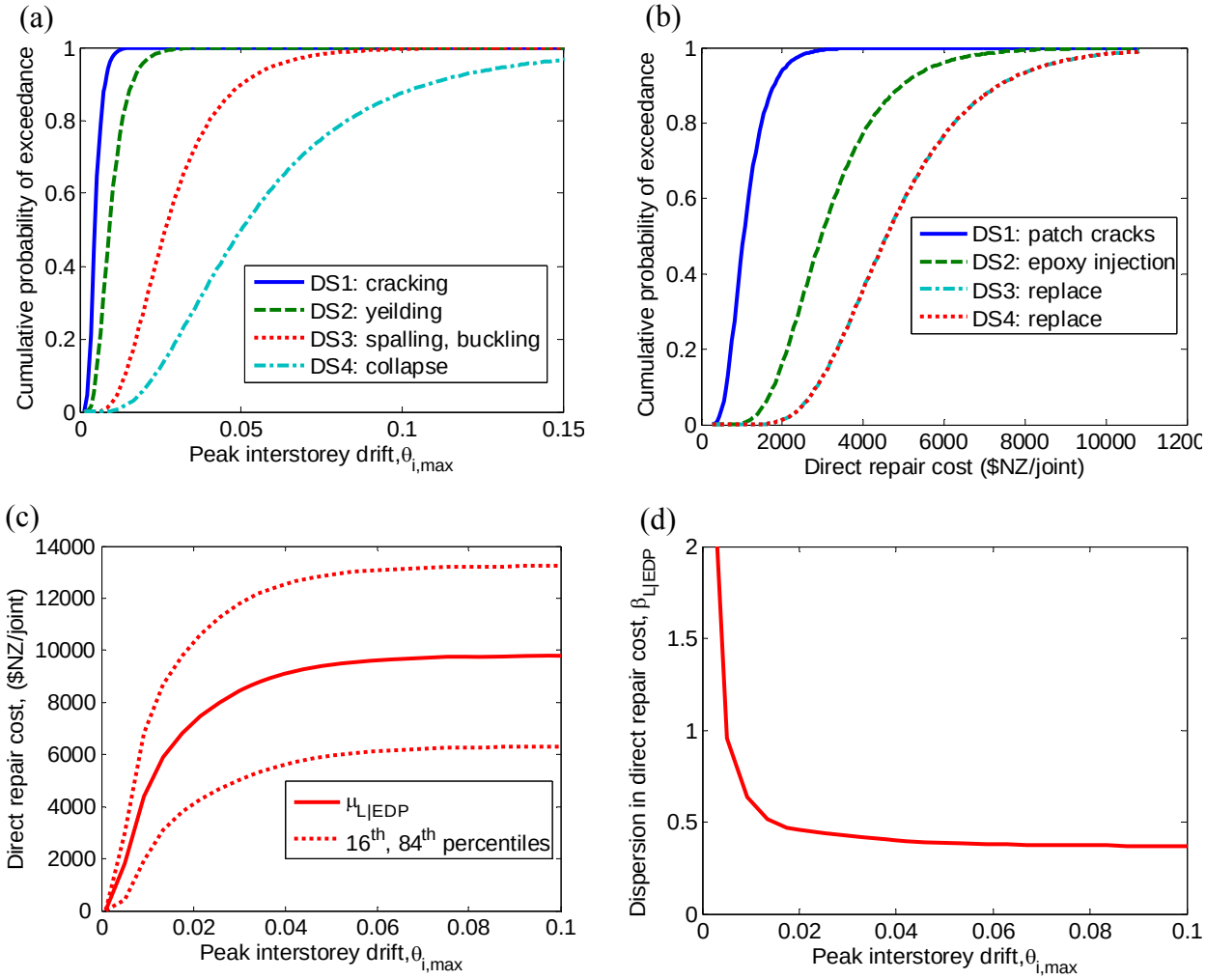


Figure 2-7: Illustration of: (a) fragility functions; (b) loss functions; (c) mean and ± one standard deviation loss|EDP; and (d) dispersion in loss|EDP for a ductile RC beam-column joint.

2.10 Loss given intensity relationship

2.10.1 *L*|*IM* for an individual component

The *L*|*EDP* relationships previously discussed are useful in observing the seismic performance of various components, however it is more insightful to combine these *L*|*EDP* relationships with the seismic response (*EDP*|*IM*) relationship in order to determine the loss as a function of ground motion intensity (*L*|*IM* relationship), both for individual components, and

for the entire structure. The mean and variance in the loss for an individual component as a function of the ground motion IM can be obtained from:

$$\mu_{L_i|IM}(im) = \int \mu_{L_i|EDP}(edp) f_{EDP|IM}(edp|im) dEDP \tag{2-8}$$

$$\sigma^2_{L_i|IM}(im) = \int \left[\mu^2_{L_i|EDP}(edp) + \sigma^2_{L_i|EDP}(edp)\right] f_{EDP|IM}(edp|im) dEDP - \mu^2_{L_i|IM}(im) \tag{2-9}$$

where $\mu_{L_i|EDP}(edp)$ and $\sigma^2_{L_i|EDP}(edp)$ are obtained from Equations (2-5) and (2-6), respectively; and $f_{EDP|IM}(edp|im)$ is the probability density function (pdf) for the *EDP*|*IM* relationship, which can be obtained from

$$f_{EDP|IM}(edp|im) = -\frac{dG_{EDP|IM,NC}(edp \mid im)}{dEDP} \tag{2-10}$$

where $G_{EDP|IM,NC}(edp \mid im)$ is obtained from seismic response results after removing global collapse cases (Figure 2-3). Hence, the *L*|*IM* relationship for a single component does not consider collapse cases and thus for brevity is not explicitly given in some mathematical notations (i.e. $f_{EDP|IM}(edp|im)$, $\mu_{L_i|IM}(im)$ and $\sigma^2_{L_i|IM}(im)$).

Figure 2-8 illustrates the *L*|*IM* relationships for two different components in the case study structure. The first (Figure 2-8a) is an RC joint which was located in the second floor of the structure (e.g. *EDP*|*IM* relationship given in Figure 2-3a and 2-3c and *L*|*EDP* relationship in Figure 2-7), while the second (Figure 2-8b) is a drywall partition located in the 8th storey of the structure (e.g. *EDP*|*IM* relationship given in Figure 2-3b and 2-3d). For both components, as would be expected, the loss due to direct damage increases as the ground motion shaking increases. For spectral acceleration levels in excess of 0.5g the drywall partition is likely to be completely damaged, while the damage continues to increase in the RC joint (indicated by an increase in the mean repair cost) up to S_a levels of 0.8g. As the seismic drift demand on the 8th floor is significantly smaller than that on the lower half of the floors in the structure, a partition located on floors with higher seismic demand (for a given level of S_a) would sustain complete damage at significantly lower levels of spectral acceleration.

2.10.2 *L*|*IM* for the entire structure

The *L*|*IM* relationship for the entire structure given no-collapse can be obtained as the summation of the *L*|*IM* relationships for all of the components comprising the structure:

$$\mu_{L_T|IM,NC}(im) = \sum_{i=1}^{Nc} \mu_{L_i|IM}(im) \tag{2-11}$$

$$\sigma^2_{L_T|IM,NC}(im) = \sum_{i=1}^{Nc} \sigma^2_{L_i|IM}(im) + 2\sum_{i=1}^{Nc}\sum_{j=1}^{i-1} \rho_{L_i,L_j|IM}(im)\sigma_{L_i|IM}(im)\sigma_{L_j|IM}(im) \tag{2-12}$$

where N_C is the number of components in the structure; $\mu_{L_T|IM,NC}(im)$ and $\sigma^2_{L_T|IM,NC}(im)$ are the mean and variance in the total loss given $IM = im$ and no collapse; and $\rho_{L_i,L_j|IM}(im)$ is the correlation coefficient between the losses in various components. The correlation coefficient defines the linear dependency between the variances of different components within the structure and depends on the location within the structure (i.e. the correlation between different seismic demands within the structure) and the type of component (i.e. the correlation between damage fragility and repair actions of different components). Determination of the correlation matrix (comprised of the correlation coefficients) is a difficult task plagued by a lack of data, and is beyond the scope of this chapter where it is simply assumed that all components are uncorrelated (a lower bound on the true uncertainty in the *L*|*IM* relationship).

The *L*|*IM* relationship for the entire structure considering both collapse and non-collapse cases can be obtained from:

$$\mu_{L_T|IM}(im) = \mu_{L_T|IM,NC}(im)\left[1 - P_{C|IM}(im)\right] + \mu_{L_T|C} P_{C|IM}(im) \tag{2-13}$$

$$\begin{aligned}\sigma^2_{L_T|IM}(im) = {} & \sigma^2_{L_T|IM,NC}(im)\left[1 - P_{C|IM}(im)\right] + \sigma^2_{L_T|C} P_{C|IM}(im) \\ & + \left(\mu_{L_T|IM,NC}(im) - \mu_{L_T|C}\right)^2 P_{C|IM}(im)\left[1 - P_{C|IM}(im)\right]\end{aligned} \tag{2-14}$$

where $\mu_{L_T|C}$ and $\sigma^2_{L_T|C}$ are the mean and variance of the loss given global collapse and $\mu_{L_T|IM}(im)$ and $\sigma^2_{L_T|IM}(im)$ are the mean and variance of the total loss given $IM = im$, once conditioning on collapse (and no collapse) has been removed. Figure 2-9 illustrates the *L*|*IM* relationship for the entire structure for: (a) loss given *IM* and no collapse; (b) loss given collapse; and (c) loss given *IM* with no conditioning on collapse or no collapse. The expected loss and dispersion in the loss given collapse were computed based on the replacement cost of the structure plus an additional 8% to account for demolition and re-design [19]. The relatively small dispersion in the loss given no collapse case is due to the lower-bound assumption made here that losses between components are uncorrelated. Figure 2-9c illustrates the total loss given *IM* for the structure once both collapse and non-collapse cases

are considered. As Equation (2-13) is simply the loss due to collapse and non-collapse weighted by the probability of each of the two mutually exclusive events occurring then it is clear that the *L*|*IM* relationship is similar to the *L*|*IM*,*NC* relationship for small levels of ground motion intensity (where P(*C*|*IM*) is small), and is similar to the *L*|*C* relationship for large levels of ground motion (where P(*C*|*IM*) is new unity).

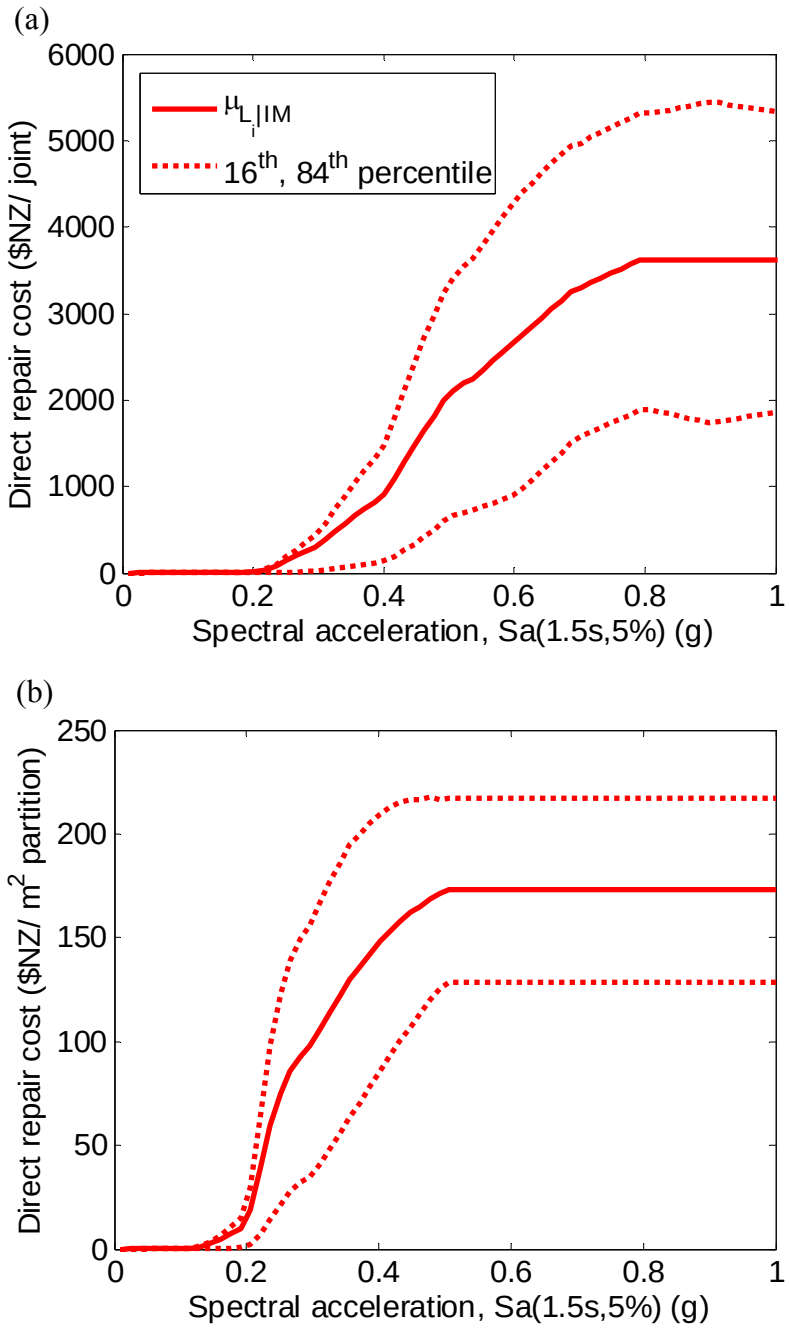


Figure 2-8: Loss given intensity relationships for (a) RC beam-column joint and (b) drywall partition.

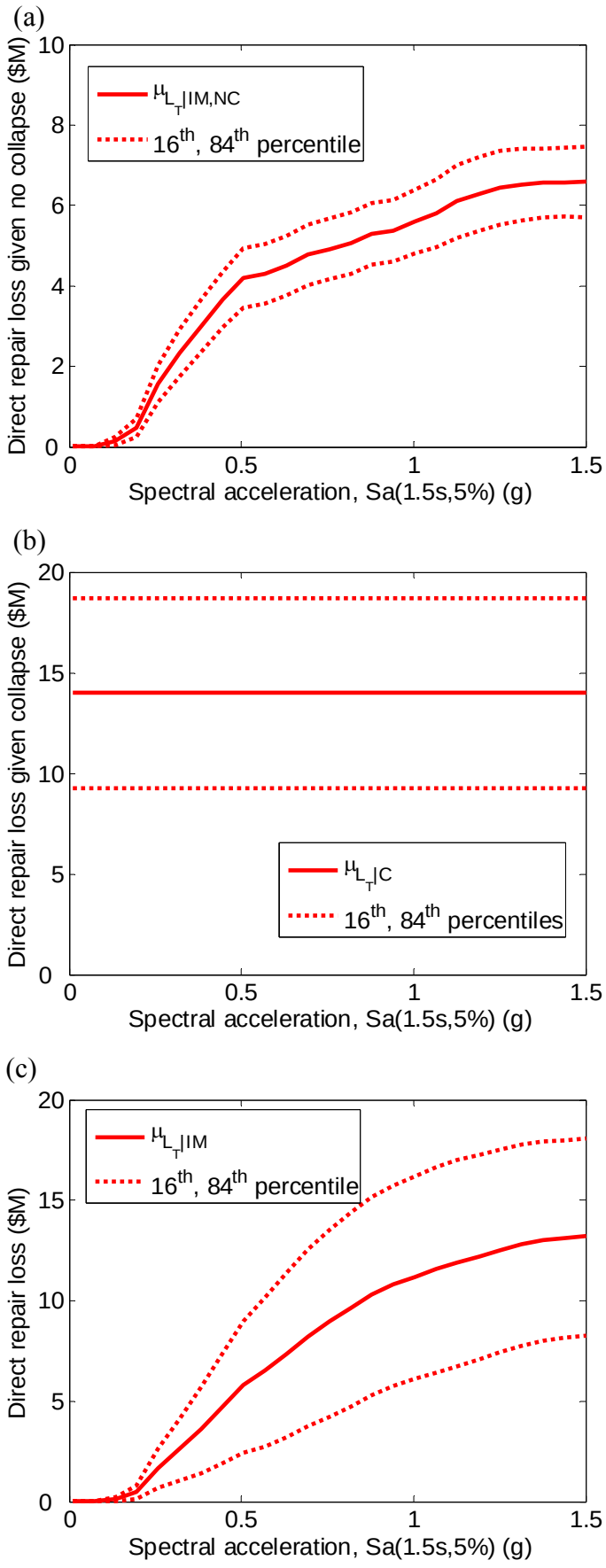


Figure 2-9: Loss given intensity for the entire structure given: (a) collapse does not occur; (b) collapse occurs; and (c) both collapse and non-collapse cases considered.

2.10.3 Deaggregation of *L|IM* by collapse

Since the total loss given IM is an accumulation of damage to many different components on various floors of the structure it is insightful to deaggregate the loss to investigate key contributors (and therefore how the loss can be reduced most effectively) [19]. Figure 2-10a illustrates the deaggregation of the total loss by collapse and non-collapse cases for the analysed structure (Equation (2-13)). As one would expect, for small levels of ground motion the probability of collapse is very small and therefore the majority of the loss is due to damage to individual components when the structure does not collapse. As the level of ground motion intensity increases the contribution of losses due to collapse increases. For example, if a ground motion observed during the 1995 Kobe earthquake occurred at the site of the structure it would be expected that 74% of the total loss in the structure would be due to the possibility of structural collapse. Figure 2-10b illustrates the effect on the *L|IM* collapse deaggregation for the same structure with a structural system assumed to be of a pre-1970's design (e.g. non-ductile, axial/shear critical RC frames). In this case it was assumed that the non-ductile structure had a collapse fragility with mean 0.5g and dispersion of 0.48. This gives an annual rate of collapse of 1.0×10^{-3} (4.9% probability of occurrence in 50 years), which is a representative collapse rate for such structures [19]. Comparing Figures 2-10a and 2-10b indicates that as the structure becomes more vulnerable to global collapse, the loss is further dominated by the collapse losses, and also the total loss for a given level of intensity also increases.

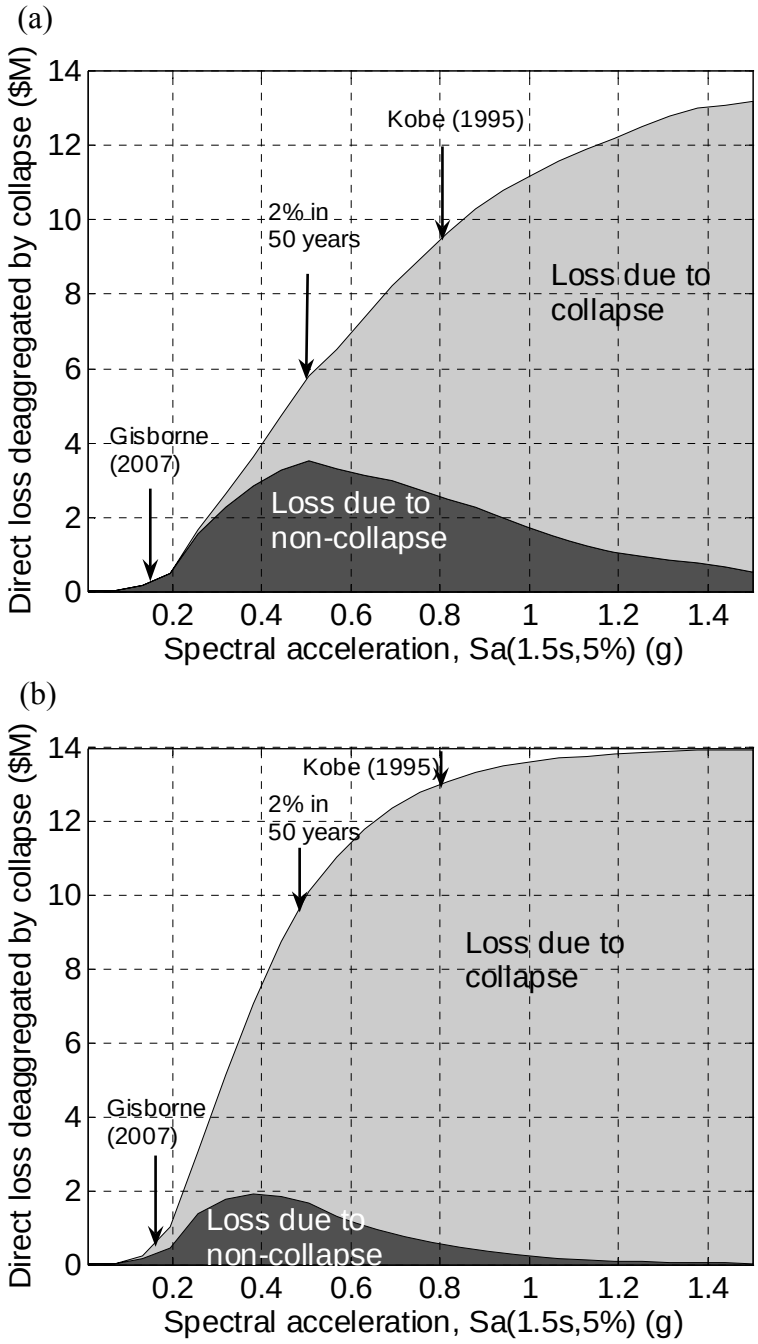


Figure 2-10: Deaggregation of the mean loss given intensity to collapse and non-collapse losses for: (a) the case study structure; and (b) for a similar structure with poor detailing.

2.10.4 Deaggregation of L|IM by component type and location in building

Figure 2-10 illustrates that for small levels of ground motion shaking a large portion of the total loss in structures is due to that which occurs in the absence of global collapse (this is particularly true for modern ductile designed structural systems as shown in the comparison of Figures 2-10a and 2-10b). Thus in the case of no-structural collapse, further insight can be obtained by deaggregation of the expected loss given no collapse (Equation (2-11)). Figure

2-11 illustrates the deaggregation of the *L|IM,NC* relationship by different component types. It is immediately apparent that over all the observed ranges of ground motion considered here, loss due to damage in structural components (i.e. RC beam-column joints and Slab-beam-column connections) make up a very minor proportion of the total loss in the structure. Non-structural components such as drywall partitions and acoustical ceilings (as well as generic components) comprise a significant proportion of the total loss. Contents such as Server and Network equipment and Computers are also significant contributors toward the total loss in the structure.

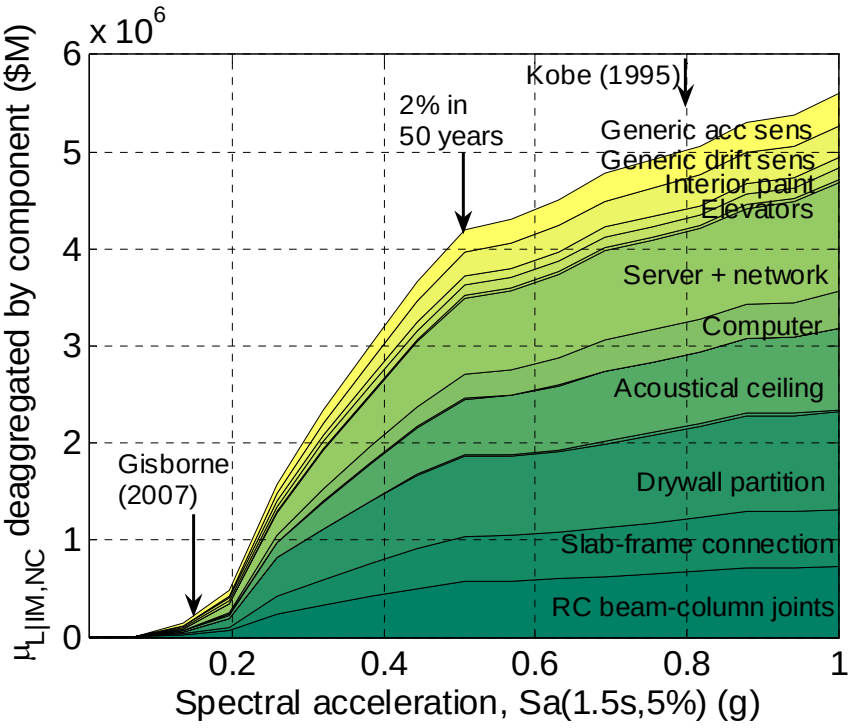


Figure 2-11: Deaggregation of the mean loss given no collapse relationship to contributions from different components.

As all of the different components within the structure have different fragilities and are subjected to different seismic demands (i.e. due to different locations in the structure) then further information can be obtained by looking at the loss deaggregation for a single level of ground motion shaking. Figure 2-12 illustrates the deaggregation of the loss given no collapse for a ground motion shaking of IM = 0.15g S_a which is approximately that observed in the recent Gisborne earthquake [28]. For this level of ground motion shaking the expected direct loss was \$0.23 M for the case study structure. Figure 2-12a illustrates that for this relatively small level of shaking the total loss is dominated by that due to non-structural and contents damage and little loss is incurred due to structural damage (24%). As noted by Mitrani-Reiser (2007) for this low level of ground shaking a large portion of the loss results from

replacement and re-painting of damaged partitions and acoustical ceilings (28%). Servers and computer equipment also have a significant contribution to the total loss (20%), and generic acceleration sensitive components (16%). Note that Figure 2-12a and 2-13a do not show components which contribute less than 1% to the total loss (e.g. exterior glazing, automatic sprinklers). Figure 2-12b illustrates the contribution of the total loss as a function of the different storey numbers. Recall that most of the components were assumed to have the same quantity for each floor in the structure, with the exception of the server and network equipment, roof mounted equipment and elevators. The server and network equipment is the primary reason for the relatively large levels of loss on the 3rd, 6th and 10th floors, while the elevator damage also makes a significant contribution to the loss on the ground floor.

Figure 2-13 illustrates the $L|IM,NC$ deaggregation for IM = 0.5g S_a which is the level of ground motion shaking with a 2% probability of exceedance in 50 years (return period approximately 2475 years) at the site. For this level of shaking the expected loss was estimated to be \$4.2 M. Trends observed in Figure 2-12a in regard to the loss being dominated by non-structural and contents damage are also evident at this higher level of ground shaking. Notable changes from Figure 2-12 however are the reduction in the proportion of damage due to re-painting of interior walls; damage to elevators, and an increase in the proportion of loss due to damage of the partitions and acoustical ceilings. The deaggregation of damage over the height of the structure is also significantly different from that at the lower level of ground motion shaking. The effect of the server and network equipment in the distribution of loss over the height of the structure is very evident, while the reduction in proportion of damage due to the elevator reduces the contribution of loss in the ground floor. Despite the significant increase in drift demands (Figure 2-4a) on the lower floors of the structure for this level of shaking relative to the results in the previous figure (S_a = 0.15g), a significant portion of the loss still originates from the upper floors. This is due to an increase in the acceleration demands on the upper floors which are of a large enough magnitude to damage a large proportion of fragile acceleration-sensitive components. Figure 2-3 illustrated that as the ground motion shaking increases, both acceleration and drift demands become largest in the lower floors of the structure (illustrated by the domination of losses in these floors in Figures 2-12a and 2-13a). This localisation of demand in the lower stories prevents components in the upper stories sustaining total damage, and thus the reason why the total loss given no collapse (Figure 2-9a) even at large IM levels, is still well below the total loss given collapse occurs.

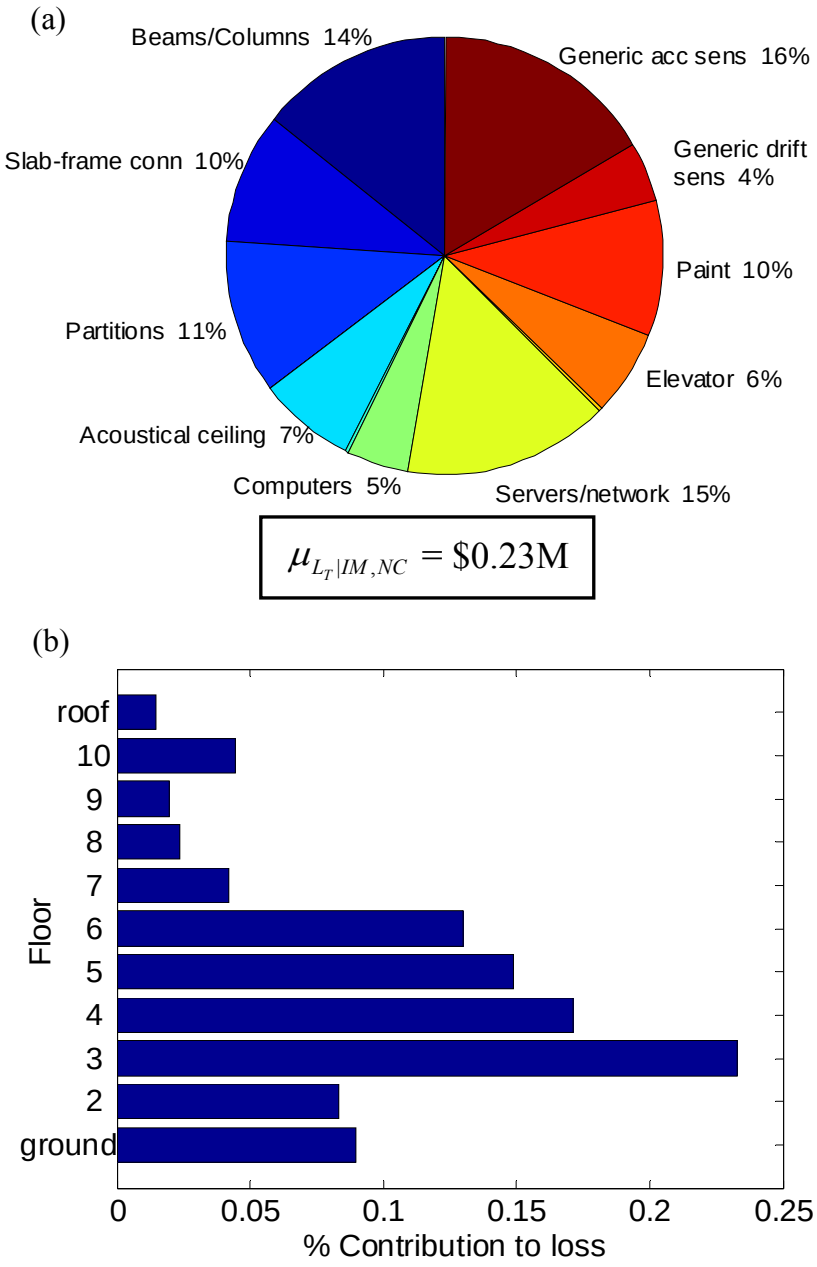


Figure 2-12: Deaggregation of the expected loss given no collapse for IM = 0.15g S_a (Gisborne, 2007) by: (a) component type and (b) by floor.

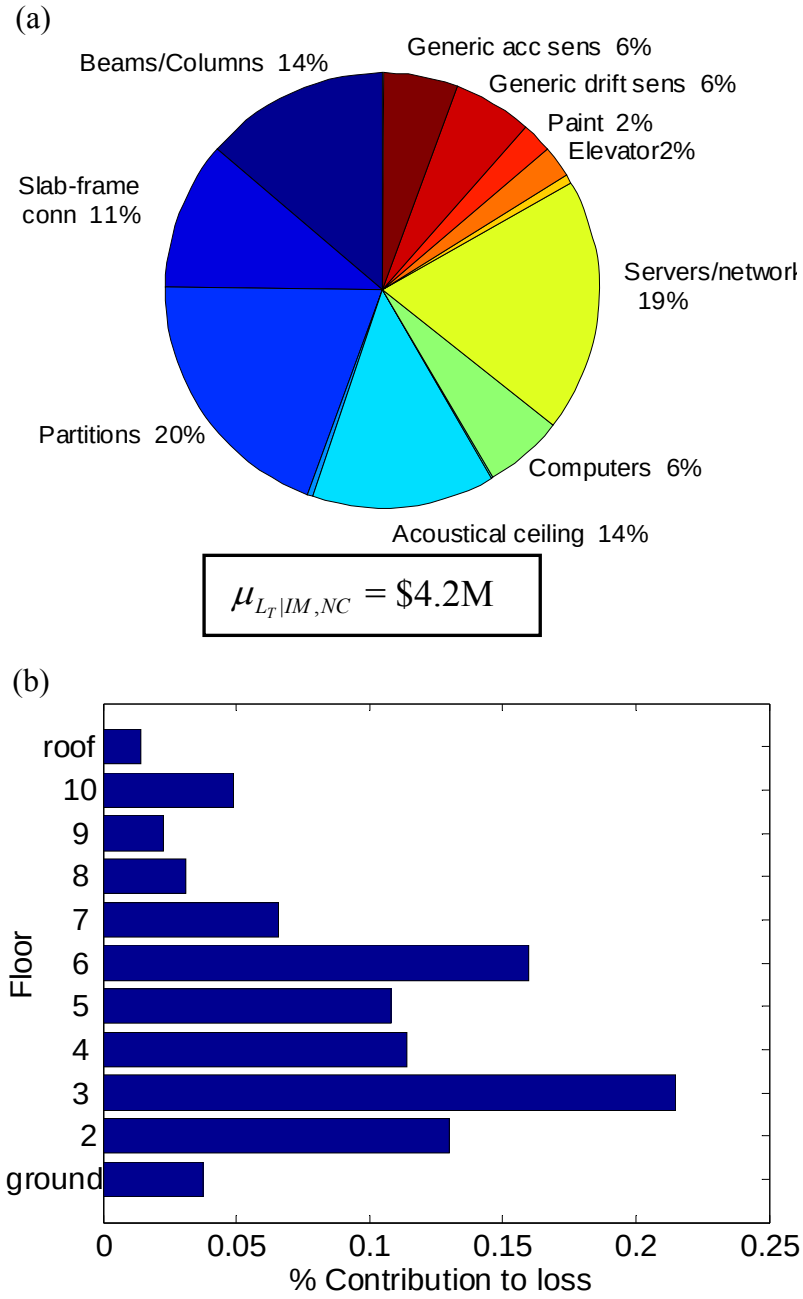


Figure 2-13: Deaggregation of the expected loss given no collapse for IM = 0.50g S_a (2% in 50 year probability of exceedance) by: (a) component type and (b) by floor.

2.11 Expected annual loss (EAL)

The above *L|IM* relationships are useful in determining the variation in loss as the ground motion increases. However, the ground motion hazard is needed in conjunction with the *L|IM* results in order to appreciate the frequency of exceedance for each of the ground motion IM levels. The *L|IM* relationship can be combined directly with the ground motion hazard to get the moments of annual loss that occurs on a per annum basis by:

$$\mu_{L_T} = \int \mu_{L_T|IM}(im) \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM \tag{2-15}$$

$$\sigma^2_{L_T} = \int \left[\sigma^2_{L_T|IM}(im) + \mu^2_{L_T|IM}(im) \right] \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM - \mu^2_{L_T} \tag{2-16}$$

where μ_{L_T} and $\sigma^2_{L_T}$ are the mean and variance of the annual loss.

Expected annual loss (EAL = μ_{L_T}) is a seismic performance measure which is particularly useful for decision makers as it contains information on the seismic performance of a structure over a range of different levels of ground motion intensity within a single number. EAL has been used extensively by Porter and co-workers [e.g. 23, 29] as well as others in risk management decision making. For the case study structure in was found that the EAL was \$11,700 which relates to approximately 0.08% of the replacement cost of the structure. Figure 2-14a illustrates the net present value (NPV) of the expected loss over time for the structure based on a discount rate of 6%. The discount rate is used to account for the time value of money and represents the difference between interest and inflation rates. Figure 2-14b illustrates the deaggregation of the EAL as a function of ground motion intensity. It indicates that the majority of the EAL is attributed to the occurrence of ground motions between 0.25-0.75g S_a (with smaller ground motions not causing significant damage, and larger ground motions occurring very infrequently). Although not presented here there are several other possible ways to deaggregate the EAL such as by components, collapse, and location in the structure (as for the *L|IM* deaggregations shown in Figures 2-10-2-13).

As an indication of the seismic performance of non code-complying structures, the case study structure with modified collapse capacity which was discussed in Figure 2-10b was found to have an EAL of \$20,500, with a large proportional of the EAL due to collapse events. Despite the significantly poor seismic performance of such a structural system, the EAL is still only 75% larger than that of the case study structure considered here. This is due to the aforementioned observation that economic loss in structures during small to moderate ground motions is dominated by damage to non-structural components and contents rather than structural components.

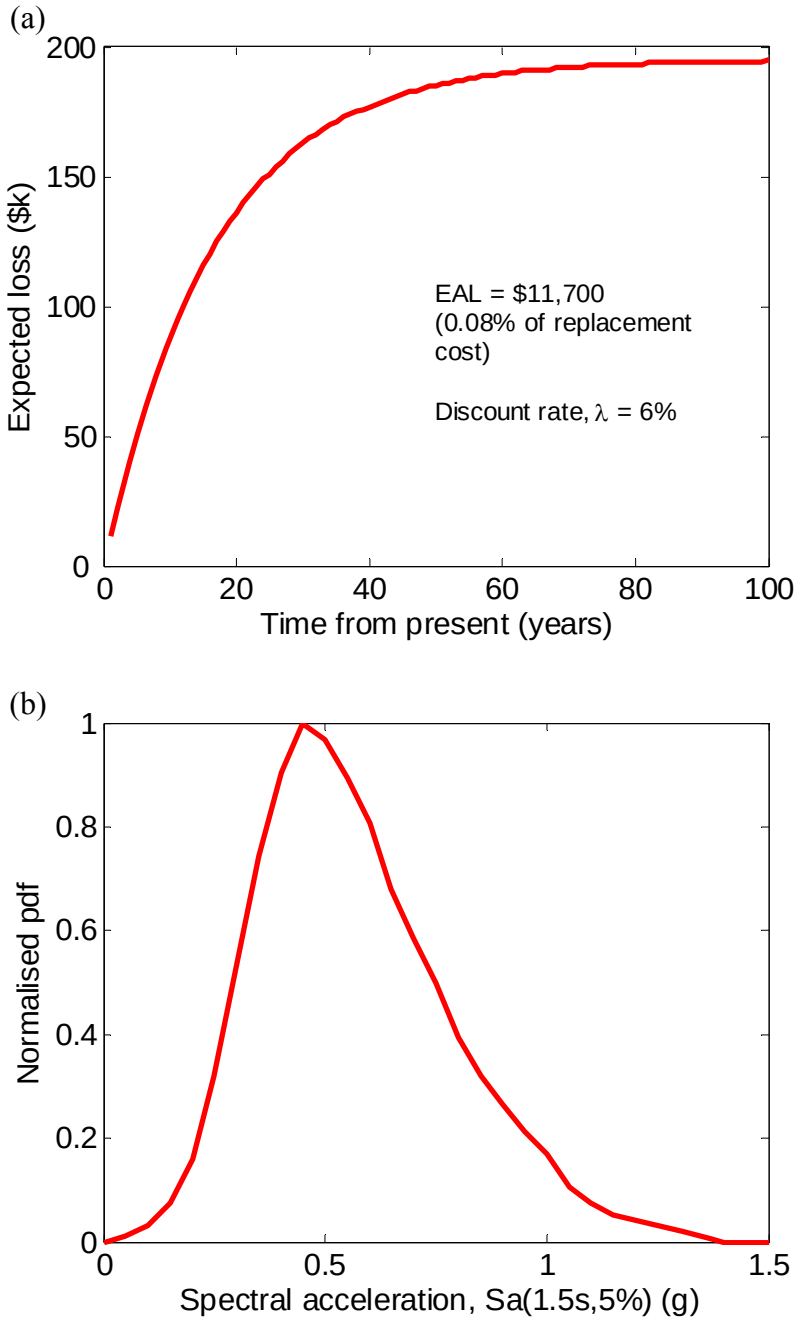


Figure 2-14: Expected loss results for case study structure: (a) over time considering net discount rate; and (b) deaggregation of EAL by intensity measure.

2.11.1 Application of EAL for retrofit decision making

Figure 2-15 illustrates an example of how EAL can be used as a method for determining the viability of different design alternatives or retrofit solutions. In this case the current structure is the case study structure with an EAL of $11,700. A retrofit solution based on the installation of viscous dampers is used to primarily reduce acceleration demands in the structure due to ground motion shaking and has an assumed cost of $40,000. The retrofit will

reduce the EAL to \$8,000. As the NPV of expected loss over time accounting for the discount rate is computed by:

$$E_L = \frac{(1-e^{-\lambda t})}{\lambda} EAL + C_R \tag{2-17}$$

where λ is the discount rate; t is the time in years; and C_R is the retrofit cost. Then by equating Equation (2-17) for the as-is structure ($C_R = 0$) and retrofitted structure the time after which the retrofit is economically feasible can be found by:

$$t_{cr} = \frac{-1}{\lambda} \ln\left(1 - \frac{\lambda}{(1-\alpha)} \frac{C_R}{EAL}\right) \tag{2-18}$$

where α is a parameter indicating the reduction in the EAL due to the retrofit (i.e. α = 8,000/11,700 = 0.68). Using Equation (2-18) it is found that this critical time for retrofit viability is 17 years (Figure 2-15). Thus the service life of the structure should be greater than 17 years in order for the retrofit to be beneficial. The example above is based on decision-making for a client who is risk neutral. When risk attitude is taken into account (decision makers are typically risk averse toward low probability high consequence events such as earthquakes) then measures such as the Certainty Equivalent (CE) should be used [29].

In the example presented above EAL was used as the performance measure of interest. In certain cases the specific performance measure used will depend on the perspective from which the decision making is made. For example, performance measures for an owner and an occupant will be different, with the owner principally interested in minimising damage and business downtime, while the client is interested in minimising human loss and contents damage.

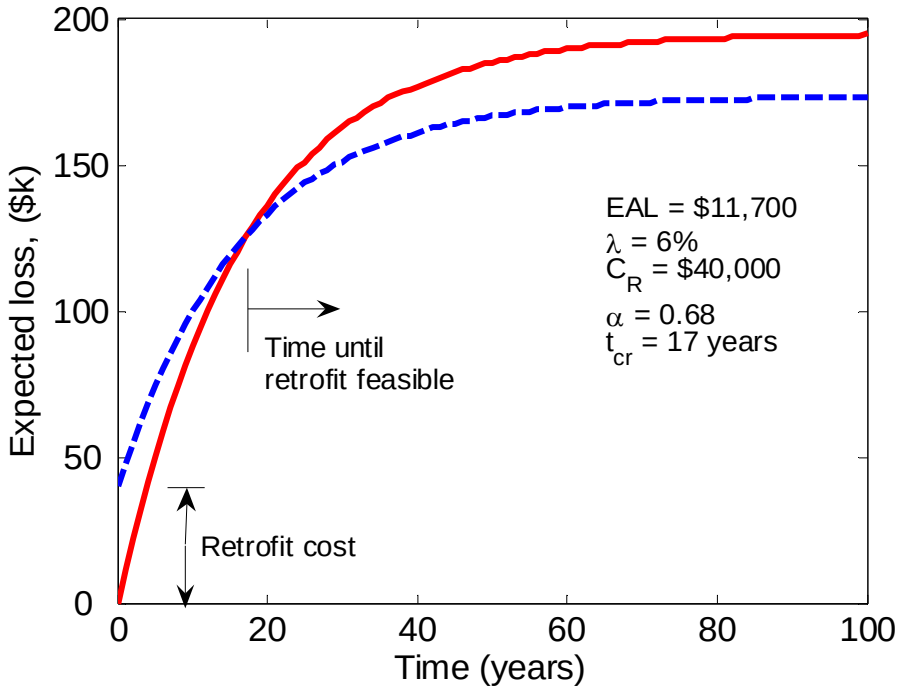


Figure 2-15: Use of EAL in determine effective retrofitting solutions.

2.12 Loss hazard

In a similar manner to the computation of the EDP hazard curves presented earlier, the L|IM relationship can be combined with the ground motion hazard curve to obtain a loss hazard curve for the structure:

$$\lambda_{L_T}(l_t) = \int G_{L_T|IM}(l_t \mid im) \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM \tag{2-19}$$

where $G_{L_T|IM}(l_t \mid im)$ is the CCDF of the $L|IM$ relationship which can be obtained by:

$$G_{L_T|IM}(l_t \mid im) = G_{L_T|IM,NC}(l_t \mid im)\left[1 - P_{C|IM}(im)\right] + G_{L_T|C}(l_t) P_{C|IM}(im) \tag{2-20}$$

where $G_{L_T|IM,NC}(l_t \mid im)$ and $G_{L_T|C}(l_t)$ are the CCDF's of the $L|IM,NC$ and $L|C$ relationships, respectively. The $L|IM,NC$ distribution is assumed to have a normal distribution based on the central limit theorem, while the $L|C$ distribution is assumed to be lognormal based on past research [19]. Figure 2-16 illustrates the loss hazard curve for the case study structure. The loss hazard curve gives the annual frequency of exceeding various levels of economic loss in the structure. For reference, the 10% and 2% probability of exceedance in 50 year curves are shown which have loss values of approximately $1.5 M and $4.5 M, respectively. The loss-hazard curve is another performance measure which can be useful in loss-based decision

making, particularly for risk-averse decision makers who are concerned with the frequency of various economic losses as opposed to just the expected annual loss.

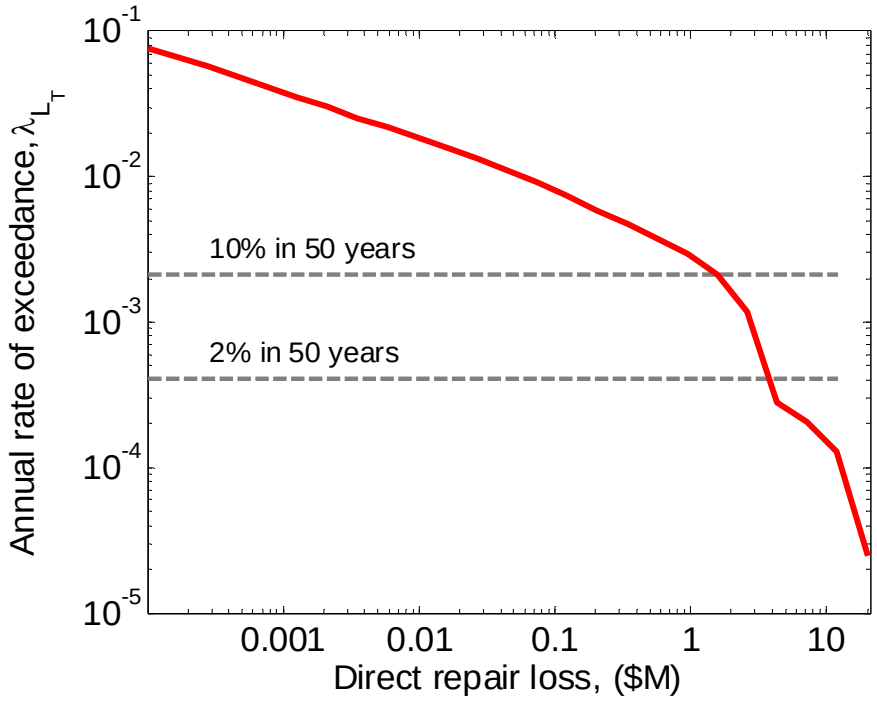


Figure 2-16: Loss hazard curve for the case study structure.

2.13 Limitations of loss assessment

Throughout this chapter various assumptions have been made regarding the loss assessment in order to focus its use in the decision making process without attending significantly to many technical details. Some of the simplifying assumptions made could cause significant error in the outputs of the loss-assessment (and the resulting decisions made) and therefore should be appropriately considered. An outline of considerations not made in the presented example is given below, along with recent research in each of these areas where appropriate.

Economic losses due to human causality and injury and business disruption/downtime were not included in the example presented in this chapter, but both have been shown to be significant in previous earthquakes (in many cases more than loss due to direct damage). Models for considering both of these losses are still in their infancy and an early attempt within this loss-assessment methodology can be found in Mitrani-Reiser, [23].

The direct repair cost estimates used in this example did not consider the effects of demand surge. Demand surge refers to the observed increase in the unit cost of labour and materials when demand for resources far exceeds the resource supply. Preliminary work on

quantifying demand surge effects can be found in Boissonnade [30].

The seismic hazard used for the site considers only the likely occurrence of ground motions due to mainshocks. It is typical following a mainshock ground motion that multiple aftershock ground motions will also occur. Yeo and Cornell [31] provide recent work illustrating how aftershock considerations can be modelled within loss estimation procedures.

Stiff soil at the site of the structure was assumed and therefore it was a valid assumption that the effects of ground motion modification due to soil non-linearities and damage to the foundation was not significant. Many structures (particularly lifelines such as bridges) are founded on pile foundations which are located on loose sands susceptible to liquefaction and lateral spreading. In such cases, ground motion modification and foundation damage may contribute significantly toward the total loss in the structure during strong ground motion shaking.

2.14 Discussion: Improving structural performance

Figures 2-12a and 2-13a illustrate that for a code-conforming structural system, economic loss due to both moderate and large ground motion shaking is mostly caused by damage to non-structural components and contents and not due to structural damage or global collapse. This indicates that the improvement of seismic performance of structural systems requires a focus on improving the performance of the entire structure and not just the structural system specifically. As an example, based on the results presented in this chapter, emerging jointed precast concrete systems [32, 33] which are designed to undergo the same levels of seismic demand as that of conventionally designed ductile structures with minimal structural damage will only mildly improve the performance of the entire system (an ambitious estimate could be made by assuming that all structural damage was avoided thus reducing the losses by approximately 25%). It should be noted however, that downtime (business disruption) losses have not been considered here and in some cases downtime due to structural damage may be excessive. Therefore significantly improved seismic performance of buildings can be obtained either by: (i) improving component fragilities via modifying connection details on non-structural components connected to the structural frame; fastening acceleration sensitive components which topple, base isolating expensive components (e.g. servers, electrical equipment); or (ii) by reducing seismic demands throughout the structure for the same level of ground shaking via increased viscous and hysteretic damping using dissipation devices or base isolation devices.

2.15 Discussion: SLAT overview

All of the loss-based computations performed in this chapter have used the specially developed software SLAT (Seismic Loss Assessment Tool) [26]. SLAT has been developed as a general-purpose loss assessment tool which is capable of performing various loss assessment procedures with various types of input data (in regard to complexity). Although SLAT is still only currently in an 'alpha' version it is intended to make the tool available in the public domain in order to allow the seismic design of structures to be based on more rational measures of seismic performance.

2.16 Conclusions

A seismic loss estimation methodology has been presented and applied for the seismic assessment of a 10 storey reinforced concrete moment frame structure. The seismic loss estimation methodology enables quantification of the seismic risk of engineered structures thus allowing consistent communication and rational decision making regarding the acceptance or mitigation of the seismic risk. It was illustrated how the seismic loss estimation methodology can be used to interpret seismic performance in terms of seismic demand and associated economic loss as a function of the ground motion intensity. Economic loss due to non-structural components and contents was shown to be significant over a large range of ground motion shaking intensities. The wealth of information regarding system performance that is possible using deaggregation within seismic loss estimation methods can be used to target areas of seismic vulnerability and therefore efficiently improve seismic performance. In particular, deaggregation illustrated that economic loss due to non-structural components and contents was significant over a large range of ground motion shaking intensities, as well as how the distribution of losses over the height of the structure can be related to seismic response of the structure. A retrofit example was used to illustrate how expected annual loss (EAL) can be used within a decision making framework to make rational loss-based decisions.

2.17 Acknowledgements

Financial support of the first author from the New Zealand Tertiary Education Commission Bright Futures scheme is appreciated. The authors also thank Mr. John Buchan of the Christchurch City Council for providing building inventory information, and Dr.

Warwick Smith of GNS for reviewing this chapter.

2.18 References

[1] Insurance Information Institute. Earthquakes: Facts and Statistics. 2006.

[2] Bull DK and Brunsdon D. Examples of Concrete Structural Design to New Zealand Standards 3101. New Zealand, 1998.

[3] Standards New Zealand. NZS 3101 1995: Part 1: Concrete Structures Standard. Wellington, NZ, 1995.

[4] Standards New Zealand. NZS 1170.5:2004. Standards New Zealand, Wellington, New Zealand., 2004.

[5] Carr AJ. Ruaumoko: Inelastic Dynamic Computer Program. Department of Civil Engineering, University of Canterbury, Christchurch, New Zealand, 2004.

[6] Stirling MW, McVerry GH, and Berryman KR. A new seismic hazard model for New Zealand. Bulletin of the Seismological Society of America 2002; 92(5): 1878–1903.

[7] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[8] Medina RA and Krawinkler H. Seismic demands for nondeteriorating frame structures and their dependence on ground motions. John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA., 2003. 373 pp.

[9] Moghadam AS and Tso WK. Pushover analysis for asymmetric and set-back multi-story buildings, in 12th World Conference on Earthquake Engineering, Auckland, New Zealand, 2000, Paper No. 1093.

[10] Fajfar P. Structural analysis in earthquake engineering – A breakthrough of simplified nonlinear methods, in 12th European Conference on Earthquake Engineering, London, UK., 2002.

[11] Penelis GG and Kappos AJ. 3D pushover analysis: The issue of torsion, in 12th European Conference on Earthquake Engineering, London, UK, 2002.

[12] Chopra AK and Goel R. A modal pushover analysis procedure to estimate seismic demands for unsymmetric-plan buildings. Earthquake Engineering and Structural Dynamics 2004; 33: 903–927.

[13] Vamvatsikos D and Cornell CA. Incremental dynamic analysis. Earthquake Engineering and Structural Dynamics 2002; 31(3): 491–514.

[14] Aslani H and Miranda E. Probability-based Seismic Response Analysis. Engineering Structures 2005; 27(8): 1151-1163.

[15] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges. Engineering Structures 2007; 29(10): 2662-2672, DOI: 10.1016/j.engstrut.2006.12.015.

[16] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[17] Jalayer F. Direct probabilistic seismic analysis: Implementing non-linear dynamic assessments. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2003, 173 pp.

[18] Shome N and Cornell CA. Structural seismic demand analysis: Consideration of "Collapse", in 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, University of Notre Dame, South Bend, Indiana, 2000, 7.

[19] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[20] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1901-1914.

[21] Elwood KJ and Moehle JP. Dynamic Collapse Analysis for a Reinforced Concrete Frame Sustaining Shear and Axial Failures. Earthquake Engineering and Structural Dynamics 2008; 37(7): 991-1012, DOI: 10.1002/eqe.787.

[22] Ibarra LF and Krawinkler H. Global collapse of frame structures under seismic excitations. University of California at Berkeley, Berkeley, CA, 2005.

[23] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[24] Buchan J. Internal report of non-structural and contents inventories for the Christchurch Council Chambers building. Christchurch, New Zealand., 2007.

[25] Taghavi S and Miranda E. Response assessment of non-structural building elements. 2005. http://peer.berkeley.edu/publications/peer_reports.html

[26] Bradley BA. User manual for SLAT: Seismic Loss Assessment Tool version 1.14. University of Canterbury Research Report No.2009-01,Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand, 2009. 94.

[27] Porter K, Kennedy R, and Bachman R. Creating Fragility Functions for Performance-based Earthquake Engineering. Earthquake Spectra 2007; 23(2): 471-489.

[28] Cousins J and McVerry G. Ground motion spectra for the 2007 Gisborne earthquake. 2007.

[29] Porter KA and Kiremidjian AS. Assembly-based vulnerability and its uses in seismic performance evaluation and risk-management decision-making. John A.Blume Earthquake Engineering Center, Stanford, CA, 2001. 216.

[30] Boissonnade A. Modelling demand surge. Workshop on Risk Acceptance and Risk Communication, Stanford University, CA, 2007.

[31] Yeo GL and Cornell CA. Stochastic Characterisation and Decision Bases under Time-Dependent Aftershock Risk in Performance-based Earthquake Engineering. John A. Blume Earthquake Engineering Centre, Stanford University., CA, 2005.

[32] Priestley MJN, Sritharan S, Conley JR, and Pampanin S. Preliminary results and conclusions from the PRESSS five-storey precast concrete test building. PCI Journal 1999; 44(6): 42–67.

[33] Mander JB and Cheng CT. Seismic resistance of bridge piers based on damage avoidance design. National Centre for Earthquake Engineering Research, Buffalo, NY, 1997.

3. Improved Seismic Hazard Model with Application to Probabilistic Seismic Demand Analysis

Bradley BA, Dhakal RP, Cubrinovski M, Mander JB, MacRae GA. Improved seismic hazard model with application to probabilistic seismic demand analysis. *Earthquake Engineering and Structural Dynamics* 2007; **36**(14): 2211-2225.

3.1 Abstract

An improved seismic hazard model for use in performance-based earthquake engineering is presented. The model is an improved approximation from the so-called 'power law' model, which is linear in log-log space. The mathematics of the model and uncertainty incorporation is briefly discussed. Various means of fitting the approximation to hazard data derived from probabilistic seismic hazard analysis (PSHA) are discussed, including the limitations of the model. Based on these 'exact' hazard data for major centers in New Zealand, the parameters for the proposed model are calibrated. To illustrate the significance of the proposed model, a performance-based assessment is conducted on a typical bridge, via probabilistic seismic demand analysis (PSDA). The new hazard model is compared to the current power law relationship to illustrate its effects on the risk assessment. The propagation of epistemic uncertainty in the seismic hazard is also considered. To allow further use of the model in conceptual calculations, a semi-analytical method is proposed to calculate the demand hazard in closed-form. For the case study shown, the resulting semi-analytical closed-form solution is shown to be significantly more accurate than the analytical closed-form solution using the power law hazard model, capturing the ‘exact’ numerical integration solution to within 7% accuracy over the entire range of exceedance rate.

3.2 Introduction

Performance-based earthquake engineering (PBEE) has emerged as a cornerstone of modern earthquake engineering as it attempts to capture the performance of structures over the full spectrum of structural behaviour, from initial elastic response through to global instability, when subjected to a range of ground-motion excitations. Seismic performance can be presented in various forms, some of the most common being: annual rate of exceeding some structural demand parameter; annual rate of exceeding some financial loss; expected annual loss (EAL); and the probability of exceeding some loss given an earthquake scenario. In order to obtain the majority of the above measures, relationships must be defined between: seismic intensity and recurrence rate; seismic intensity and structural demand; and structural demand and financial loss. The focus of this chapter, in particular, is the relationship describing the occurrence over time of a given ground-motion intensity measure. This relation, commonly in the form of a ground-motion intensity measure and annual rate of exceedance, is typically obtained by conducting probabilistic seismic hazard analysis (PSHA) [1-3]. It is advantageous to define a relationship to represent this so-called 'seismic hazard curve' so that PBEE assessments can be made using analytical or numerical integration techniques, allowing for the propagation of uncertainty in the hazard model.

Sewell *et al.* [4] proposed the following power law expression for the relationship between annual rate of exceedance and ground-motion intensity:

$$\nu\ (IM) = k_0 IM^{-k} \tag{3-1}$$

Where IM = ground-motion intensity; $\nu(IM)$ = annual rate of exceedance of a ground-motion of intensity IM; and k_0 and k are empirical constants. As seismic hazard curves are typically plotted on a log-log scale, Equation (3-1) is linear in log-log space. This form of parametric equation is primarily used when combined with a similar power law relationship for seismic intensity to demand, which permits a closed form solution for the demand hazard to be obtained (e.g. [4-8], Appendix A). This closed-form solution will provide a reasonable estimate of the demand hazard in the range of exceedance rate that the constants of Equation (3-1) are fitted. There are several methods of fitting these parameters, for example it was proposed [8] that the curve defined by Equation (3-1) be fitted through seismic hazard data at the Design Basis Earthquake (DBE) and Maximum Considered Earthquake (MCE) intensity levels which have 10% and 2% probabilities of exceedance in 50 years, respectively.

Constraining the curve to pass through these points yields the following parameter values:

$$k = \frac{\ln\left(v_{DBE} / v_{MCE}\right)}{\ln\left(IM_{MCE} / IM_{DBE}\right)} \quad , \quad k_0 = v_{DBE}\left(IM_{DBE}\right)^k \tag{3-2}$$

where IM_{DBE}, IM_{MCE}, v_{DBE}, v_{MCE} are the ground-motion intensities and annual rates of exceedance at the DBE and MCE intensity levels, respectively; and ln(.) is the natural logarithm of (.). A typical comparison of a seismic hazard curve for a Wellington (NZ) site (obtained by performing PSHA) and Equation (3-1) is given in Figure 3-1. It can be seen that due to the typical 'concave from below' shape of the hazard curve, Equation (3-1) significantly over estimates the hazard for ground-motion intensities below the DBE and above the MCE intensity levels, respectively. Equation (3-1) also slightly underestimates the hazard for intensities between the DBE and MCE levels. For example, at IM = 0.1g *Sa*, Equation (3-1) over predicts the rate of exceedance by a factor of eight. Hence, while the power law is an adequate local approximation to the hazard curve in the vicinity of the DBE and MCE , over several orders of magnitude the rate of exceedance, v , is poorly approximated (i.e. from v = 10^0-10^{-6}).

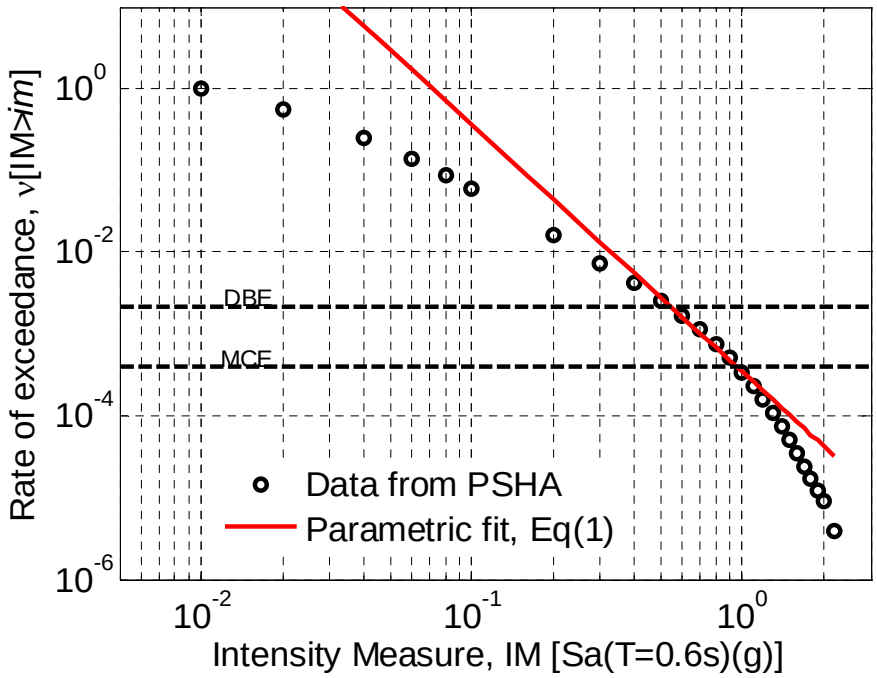


Figure 3-1: Comparison on hazard data from PSHA fitted by Equation (3-1).

This inadequacy of Equation (3-1) is not significant for performance-based design if

only a small range of exceedance frequencies are considered (and hence Equation (3-1) can be fitted appropriately for this range). However, for financial risk calculations, where the full range of the hazard is of interest the use of the power law model will lead to significant inaccuracies. Previous researchers [9] have tried to alleviate this inaccuracy for more frequent earthquake events by only considering rates of exceedance up to a certain threshold value when using the power law model. The value of this threshold is however subjective and consequently not applicable in general. Others [10] have tried fitting a lognormal distribution to the data, but its implementation is difficult as it cannot consider rates greater than one and its parametric form involves the error function, erf(**.**).

While it is possible to perform PBEE calculations using the raw data from the seismic hazard curve directly, for certain sites the hazard data provided is sparse and significant interpolation between the data points is required; e.g. Kunnath *et al*. [11] used Equation (3-1) to interpolate between three hazard data points.

Therefore, it can be seen that a parametric curve which is non-linear in log-log space and more accurately captures the actual hazard data is required for use in financial risk assessments. This chapter aims at developing improved parametric seismic hazard curves based on the above objectives. A semi-analytical closed form solution for the demand hazard using the new hazard model is presented, allowing the demand hazard to be computed without requiring numerical integration. This allows the use of the proposed hazard model in 'rapid' calculations of the demand hazard, similar to the analytical solution that can be obtained utilizing Equation (3-1), but with significantly enhanced accuracy over a large range of demand.

3.3 Hyperbolic model in log-log space

3.3.1 Model development

As the shape of the hazard curves typically have a 'concave from below' global shape (in log-log space), then it would seem reasonable to approximate the curve with a hyperbola of the form $y = \alpha/x$. Figure 3-2 illustrates the use of a reference origin that can be used to envisage how the hyperbola can be expressed in the $ln(v)$–$ln(IM)$ plane. The parametric curve has both vertical and horizontal asymptotes and is given by:

$$\ln(\nu) - \ln(\nu_{asy}) = \frac{\alpha}{\ln(IM) - \ln(IM_{asy})} + \varepsilon \tag{3-3}$$

where ν_{asy} and IM_{asy} are the horizontal and vertical asymptotes, respectively; α is constant; and ε is a random variable with mean zero and standard deviation β_H. Hence by rearranging, Equation (3-3) can be expressed as either a function of ν or IM, the median values of which are given in Equation (3-4). The three unknown parameters ν_{asy}, IM_{asy}, and α are determined using data fitting techniques, as described in the following sub-section.

$$\hat{\nu} = \nu_{asy} \exp\left[\alpha\left\{\ln\left(\frac{IM}{IM_{asy}}\right)\right\}^{-1}\right]$$
$$\widetilde{IM} = IM_{asy} \exp\left[\alpha\left\{\ln\left(\frac{\nu}{\nu_{asy}}\right)\right\}^{-1}\right] \tag{3-4}$$

The random variable, ε (Equation (3-3)), can be used to account for epistemic (modelling) uncertainty in the seismic hazard. This epistemic uncertainty is obtained through the use of logic tree weightings for different ground-motion prediction relationships [12-14].

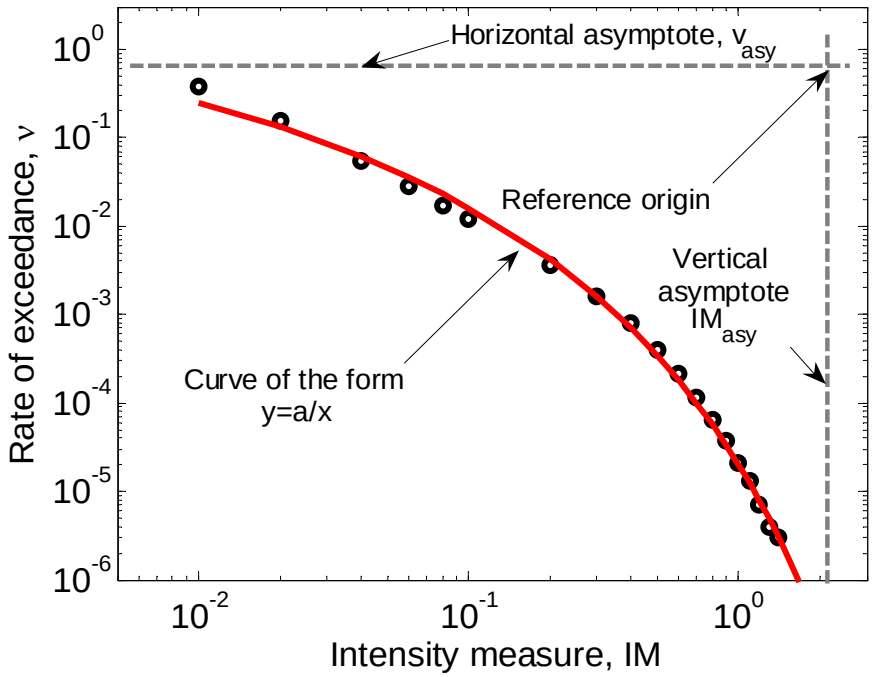


Figure 3-2: Concept of hyperbolic curve fitted to hazard data.

3.3.2 Fitting to PSHA data

To determine the parameters of the proposed hazard equation for given hazard data, the

technique of (non-linear) least-squares regression was used. In the following discussions the hyperbolic model is used in the first form of those given in Equation (3-4), so that errors are measured as deviations of v between the data and the model. As the overall shape of the hazard curve is of interest then it is desired to minimise the relative error between the data and the proposed curve and not the absolute error. The latter would lead to very accurate prediction of the data with large values of v, but poor prediction of small values. Equivalently, it is typical to minimise the logarithms of the error; then the least squares problem becomes:

$$\text{Minimise } R = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} \left[\ln(v_i) - \ln(v(IM_i))\right]^2 \tag{3-5}$$

where v_i = data points obtained via PSHA; $v(IM_i)$ = value of v obtained from parametric equation; and r_i = the least square residual for each data point.

A measure of the 'goodness-of-fit' of the parametric curve to the seismic hazard data can be objectively determined from the standard deviation (denoted as β_F) of the residuals, r_i, obtained from the regression analysis. The lower the value of β_F, the better the hyperbolic model fits the raw hazard data. Table 3-1 (discussed in the following section) gives the values of β_F for several regions in New Zealand.

3.4 Application to seismicity data

To illustrate the applicability of the proposed hyperbolic model, seismic hazard curve data for the main centres in New Zealand was obtained from Stirling *et al.* [15]. When the least squares regression is performed for both Peak Ground Acceleration (PGA) and elastic Spectral Acceleration (S_a) at a period of 1.5 seconds then Figures 3-3a and 3-3b result. It can be seen that the accuracy of the hyperbolic model is maintained over the full range of data for both regions of high and low seismicity. Only the *PGA* seismic hazard curve for Dunedin is poorly approximated by the parametric curve due to its large 'curvature' for large *IM* values and then smaller curvature at lower *IM* values (here 'curvature' refers to the second derivative of the curve in log-log space). In this case, it was selected to perform the regression on the data corresponding to the higher values of *IM*. This was selected because larger *IM* values will likely cause more structural damage and therefore have more engineering significance. Hence, for the Dunedin hazard data the first three data points were removed from the least-squares regression. The values of the three parameters for each of the *PGA* seismic hazard

curves in Figure 3-3 and the associated standard deviations from the regression analyses are presented in Table 3-1.

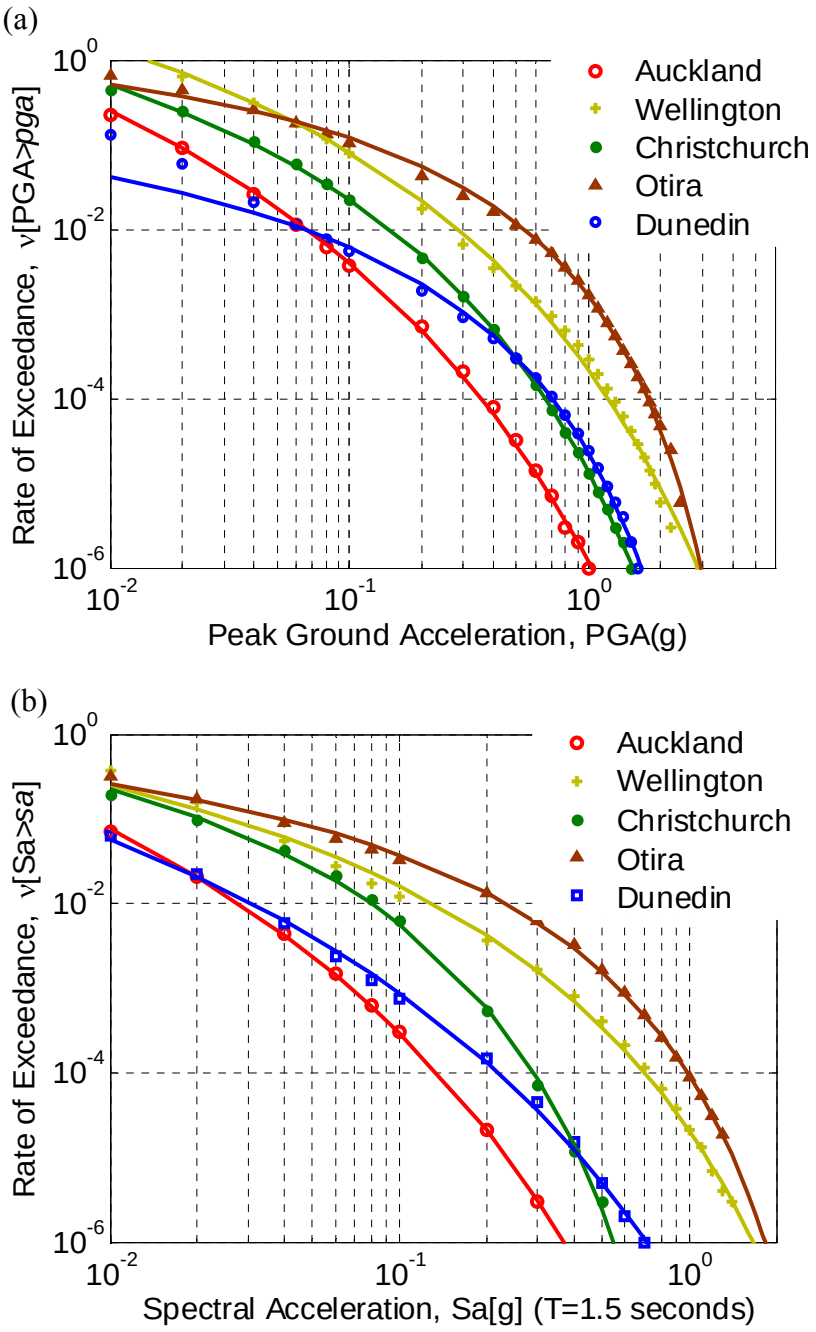


Figure 3-3: Fitted seismic hazard *PGA* and S_a data for New Zealand: (a) seismic hazard data for *PGA* fitted using Equation (3-4); and (b) seismic hazard data for *Sa*(T =1.5 s) fitted using Equation (3-4).

As with any curve fitting of data, the primary limitation of the parametric curve given by Equation (3-4) is its use in extrapolation. Asymptotes on the maximum rate of exceedance

and ground-motion intensity are requirements based on physical principles. The parametric relationship proposed has both horizontal and vertical asymptotes. However, because the parameters of the relationship are determined based on the data points within a specific range, the values of the asymptotes may not be consistent with those of different regions. Overall the range of hazard up to return periods of one million years ($v = 1x10^{-6}$) would be considered as sufficient to use for the assessment (in particular, calculating EAL which requires integration over the full range of *IM*), and therefore in the opinion of the authors no extrapolation of the parametric curve is required to obtain suitably accurate results when conducting performance-based assessments.

Table 3-1: Hazard curve parameters for various regions to be used in Equation (3-4) for *PGA*.

Region	v_{asy}	IM_{asy}	α	β_F
Auckland	98450	126	121.6	0.12
Wellington	6617	81.7	75.9	0.20
Christchurch	1221	29.8	62.2	0.06
Otira	9.95	10.5	20.5	0.14
Dunedin	1.8	10.3	26.3	0.13

3.5 Application to probabilistic seismic demand analysis

In the following two sections the propagation of the effects of the seismic hazard curve is investigated through probabilistic seismic demand analysis (PSDA) by computing the drift hazard curve for a typical bridge pier designed to New Zealand standards [16].

3.5.1 Bridge details

The prototype bridge pier is 7m high and taken from a typical 'long' multi-span highway bridge on firm soil with 40m longitudinal spans and a 10m transverse width. The seismic weight of the superstructure was calculated to be 7000 kN. The bridge was assumed to be located in Wellington, New Zealand, The fundamental period of the pier was 0.6 seconds. A computational model of the bridge pier was constructed using the nonlinear finite element program Ruaumoko [17]. The pier was modeled as a Single-Degree of Freedom (SDF), with a modified Takeda hysteresis model [18] for the force-displacement response of the pier column, and 5% viscous damping was assumed. The computation model was calibrated based on experimental results [19].

Further design details and experimental modelling of the pier can be found elsewhere [19, 20].

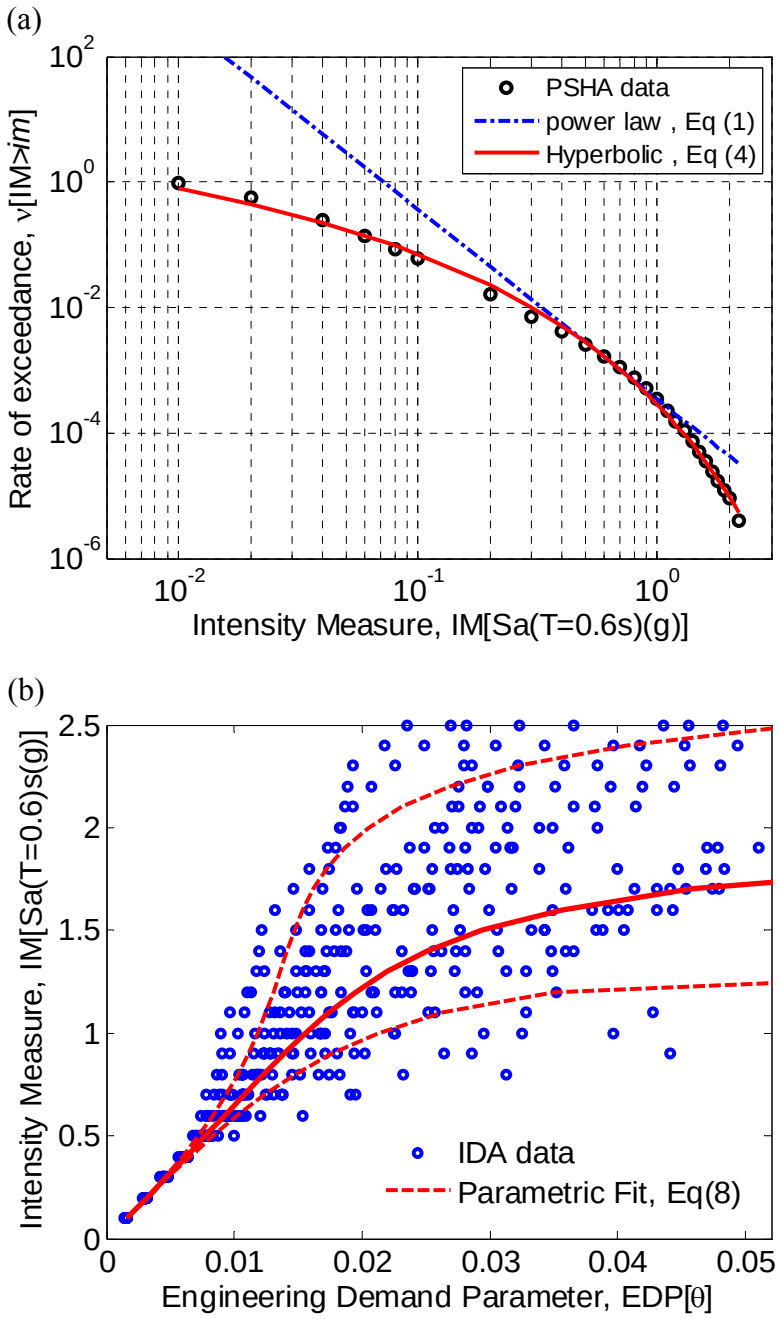


Figure 3-4: (a) Hazard model; and (b) IDA demand model curves.

3.5.2 Site seismic hazard

Seismic hazard data for the site was obtained from Stirling *et al*. [15]. The *IM* selected was the elastic spectral acceleration at the fundamental period of the structure, as it typically gives rise to a low dispersion ('efficiency') in the structural demand-response (*EDP*|*IM*) relationship, and is generally independent of magnitude and source-distance for carefully

selected ground-motion records ('sufficiency') [7, 21, 22]. From the hazard data, both power law (Equation (3-1)) and hyperbolic (Equation (3-4)) parametric equations were fitted to the data, as shown in Figure 3-4a.

3.5.3 Structural response analysis

Due to the lack of large earthquakes in the Wellington region over the past 100 years, despite its known high seismicity, there are insufficient regional ground-motion records to carry out a performance-based assessment. Therefore a suite of ground-motion records, previously used by Vamvatsikos and Cornell [23] were adopted, and are presented in Table 3-2. These records, which were all recorded on firm soil, have magnitude and distance ranges of 6.5-6.9 and 15.1-31.7 km, respectively.

Table 3-2: Ground motion records used in seismic response analysis.

No	Event	Year	Station	ϕ^{*1}	M^{*2}	R^{*3} (km)	PGA (g)
1	Loma Prieta	1989	Agnews State Hospital	90	6.9	28.2	0.159
2	Imperial Valley	1979	Plaster City	135	6.5	31.7	0.057
3	Loma Prieta	1989	Hollister Diff. Array	255	6.9	25.8	0.279
4	Loma Prieta	1989	Anderson Dam	270	6.9	21.4	0.244
5	Loma Prieta	1989	Coyote Lake Dam	285	6.5	22.3	0.179
6	Imperial Valley	1979	Cucapah	85	6.9	23.6	0.309
7	Loma Prieta	1989	Sunnyvale Colton Ave	270	6.9	28.8	0.207
8	Imperial Valley	1979	El Centro Array #13	140	6.5	21.9	0.117
9	Imperial Valley	1979	Westmoreland Fire Sta.	90	6.5	15.1	0.074
10	Loma Prieta	1989	Hollister South & Pine	0	6.9	28.8	0.371
11	Loma Prieta	1989	Sunnyvale Colton Ave	360	6.9	28.8	0.209
12	Superstition Hills	1987	Wildlife Liquefaction Array	90	6.7	24.4	0.180
13	Imperial Valley	1979	Chihuahua	282	6.5	28.7	0.254
14	Imperial Valley	1979	El Centro Array #13	230	6.5	21.9	0.139
15	Imperial Valley	1979	Westmoreland Fire Sta.	180	6.5	15.1	0.110
16	Loma Prieta	1989	WAHO	0	6.9	16.9	0.370
17	Superstition Hills	1987	Wildlife Liquefaction Array	360	6.7	24.4	0.200
18	Imperial Valley	1979	Plaster City	45	6.5	31.7	0.042
19	Loma Prieta	1989	Hollister Diff. Array	165	6.9	25.8	0.269
20	Loma Prieta	1989	WAHO	90	6.9	16.9	0.638

[1] Component
[2] Moment Magnitude
[3] Closest distance to fault rupture surface
Source: PEER Strong Motion Database, http://peer.berkeley.edu/smcat/

Using the computational model of the bridge pier, Incremental Dynamic Analysis (IDA) [24] was conducted to generate the data to characterize the conditional *EDP*|*IM* relationship. The IDA was carried out using the elastic spectral acceleration at the fundamental period of vibration as the intensity measure (IM), and the maximum deck drift as the engineering demand parameter (EDP). The resulting IDA data from the structural analyses is presented in Figure 3-4b. The conditional *EDP*|*IM* relationship was then parameterized using Equation (3-6) developed by Shome and Cornell [25], which is based on separating the mutually exclusive and collectively exhaustive cases of structural collapse and non-collapse, and a power law expression to describe the *EDP*|*IM* relation given collapse does not occur.

$$P_{EDP|IM}(edp|im) = P_{EDP|IM,NC}(edp|im)\left[1 - P_{C|IM}(im)\right] + P_{C|IM}(im) \tag{3-6}$$

where $P_{C|IM}(im)$ = the probability of collapse given $IM = im$; and $P_{EDP|IM,NC}(edp|im)$ is the probability of $EDP > edp$ given $IM = im$ and calculated assuming *EDP* given *IM* (*EDP*|*IM*) is lognormally distributed with logarithmic mean given by Equation (3-7), and lognormal standard deviation (dispersion) $\beta_{lnEDP|IM}$.

$$\mu_{\ln EDP|IM}(im) = \ln(a) + b.\ln(im) \tag{3-7}$$

where *a* and *b* are empirical constants determined by regression on the IDA data.

The 16th, 50th and 84th percentile curves (16th and 84th percentiles are one standard deviation from the median) are shown on Figure 3-4b. The bridge was deemed to collapse at a drift of 4% due to significant P-Δ effects from the superstructure. The aleatoric uncertainty, $\beta_{lnEDP|IM}$, in the *EDP* for a given *IM* was modelled using a hyperbolic tangent function, while the variation of the collapse probability with *IM* was assumed to follow a lognormal distribution [21]. A comparison of the parametric fits for dispersion and collapse probability and the raw data points are presented elsewhere [26].

3.5.4 Displacement hazard

Using both the seismic hazard and IDA parametric curves the displacement hazard of the pier can be obtained using the convolution integral presented by Deierlein *et al*. [27]:

$$\nu_{EDP}(edp) = \int P_{EDP|IM}(edp\,|\,im)\left|d\nu_{IM}(im)\right| \tag{3-8}$$

where the integration is over a range of *IM* values which have significant influence on the solution.

Equation (3-8) is then computed using numerical integration, the results of which are presented in Figure 3-5. It can be seen that in the immediate region surrounding the DBE and MCE levels the drift hazard curves given by the hyperbolic model and the power law model are similar. This is as to be expected considering the power law curve is fitted through the DBE and MCE data points. However, as expected the power law relationship significantly over-predicts the drift hazard in the region of $v > v(\text{DBE})$. While the power law relationship also over-predicts the EDP for more intense ground-motions ($v < 5\text{x}10^{-4}$), it is not as significant as would be expected based on the shape of the seismic hazard curves. The reason for this can be attributed to the fact that for these more intense ground-motions, the *EDP* exceeds the drift representing structural collapse, which is illustrated by the 'flattening' of the drift hazard curves around $v \sim 2\text{x}10^{-4}$. Therefore, it can be said that the extent of over-prediction of the power law relationship in the region of large ground-motion intensities is dependent on the seismic capacity of the structure.

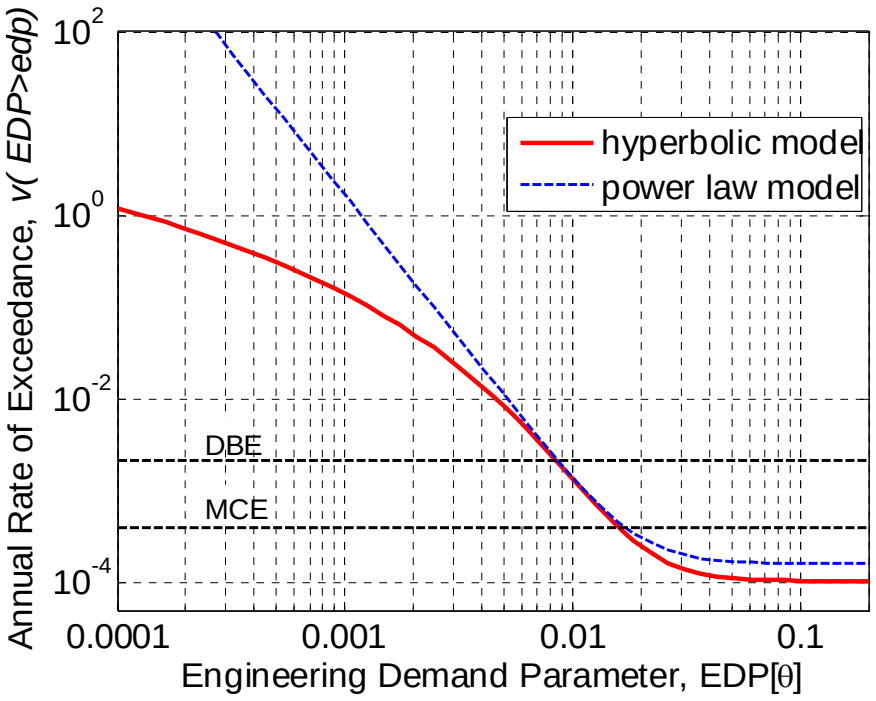


Figure 3-5: Drift hazard curves computed using power-law and hyperbolic hazard models.

3.6 A semi-analytical closed-form solution for annual frequency of demand

The attractiveness of the power law model is that it can be used to obtain a closed form solution for the frequency of exceedance of demand (herein referred to as drift hazard) for a

given structure at the site of interest (Appendix A). The mathematical form of the hyperbolic model does not directly permit such a solution. However, by inspection and appropriate modification of the closed form solution utilizing the power law hazard model, it is possible to obtain a semi-analytical solution for the drift hazard using the proposed hyperbolic model. Details of this semi-analytical solution are presented in this section. Note that unlike Figure 3-5 in the previous section (which uses Equations (3-6)-(3-8)), the closed form solution discussed in this section does not consider the onset of structural collapse. The potential incorporation of structural collapse is addressed in the discussion.

3.6.1 Mathematical details

It has been shown [4-8] that by using power law relationships for the median seismic hazard (Equation (3-1)) and structural response (Equation (3-7)) relationships, and assuming the *EDP|IM* relationship is lognormally distributed with constant dispersion ($\beta_{EDP|IM}$), a closed form solution for the demand hazard can be obtained, as given by Equation (3-9):

$$\nu(edp) = k_0 \left(\frac{edp}{a}\right)^{-\frac{k}{b}} \exp\left(\frac{1}{2}\frac{k^2}{b^2}\beta^2_{EDP|IM}\right) \tag{3-9}$$

where exp(•) is the exponential function. Jalayer [8] then suggested simplifying the presentation of Equation (3-9) by introducing the idea of the "*im* that corresponds to *edp*" defined as:

$$IM^{edp} = \left(\frac{edp}{a}\right)^{\frac{1}{b}} \tag{3-10}$$

Equation (3-10) is also the solution of Equation (3-7) for a given *edp* value. This can be interpreted as the *IM* value corresponding to the *EDP*, from the median *EDP|IM* relationship. This then allows Equation (3-9) to be expressed as follows:

$$\nu(edp) = k_0 \left(IM^{edp}\right)^{-k} \exp\left(\frac{1}{2}\frac{k^2}{b^2}\beta^2_{EDP|IM}\right) = \nu\left(IM^{edp}\right)\exp\left(\frac{1}{2}\frac{k^2}{b^2}\beta^2_{EDP|IM}\right) \tag{3-11}$$

Hence, from the right hand side of Equation (3-11), the demand hazard can be viewed as the value obtained from the median seismic hazard (Equation (3-1)) and median demand (Equation (3-7)) relations and an exponential term, which represents the increase in the hazard due to uncertainty. In particular the exponential term is a function of *k*.

Firstly, the 'median demand hazard' (defined as $\nu(IM^{edp})$ in Equation (3-11)) for a given

EDP value, can be determined by using the hyperbolic model by computing IM^{edp} from Equation (3-10) and then using Equation (3-4) with $IM = IM^{edp}$.

Secondly, it should be noted that geometrically k is simply a local approximation of the gradient of the seismic hazard curve in log-log space. The only difference between the power law and hyperbolic hazard models in this regard is that k is constant for the power law model, but not for the hyperbolic model. Hence, by differentiating Equation (3-3) with respect to $ln(IM)$, the equivalent local gradient of the hyperbolic model is given by:

$$k_{eq} = \frac{\partial[\ln(v(IM))]}{\partial[\ln(IM)]} = \frac{-\alpha}{\left[\ln\left(IM \big/ IM_{asy}\right)\right]^2} \tag{3-12}$$

where k_{eq} = equivalent local gradient of the seismic hazard curve in log-log space, and the value of IM to be used is discussed below.

It would seem intuitive that the IM value to be used (in Equation (3-12)) would be IM^{edp}. Figure 3-6 presents the values of the drift hazard curve obtained using $IM = IM^{edp}$ to compute k_{eq} in Equation (3-12), compared to the computed 'median demand hazard' ($v(IM^{edp})$), and the 'exact demand hazard' (which is obtained by direct numerical integration of Equation (3-8)). It is immediately obvious that the exponential term (the factor in Equation (3-11) giving rise to the difference between the median demand hazard curve and the exact demand hazard) is too large. This occurs because for a given value of *EDP*, the median ground-motion hazard value is always smaller than the exact ground motion hazard value. From Figure 3-3 it can be seen that the slope of the hyperbolic hazard curve increases (in absolute terms) as $v(IM>im)$ decreases. Therefore using $IM = IM^{edp}$ gives an unreasonably large k_{eq}, which is further amplified within the exponential term. Use of the mean (as opposed to median) demand curve does not resolve this problem. Therefore k_{eq} should be obtained using Equation (3-4) as before, but should be the IM corresponding to the 'exact demand hazard' value, as opposed to the 'median demand hazard' value. As the 'exact demand hazard' value is not actually known, implementing the equation in this form requires an iterative process.

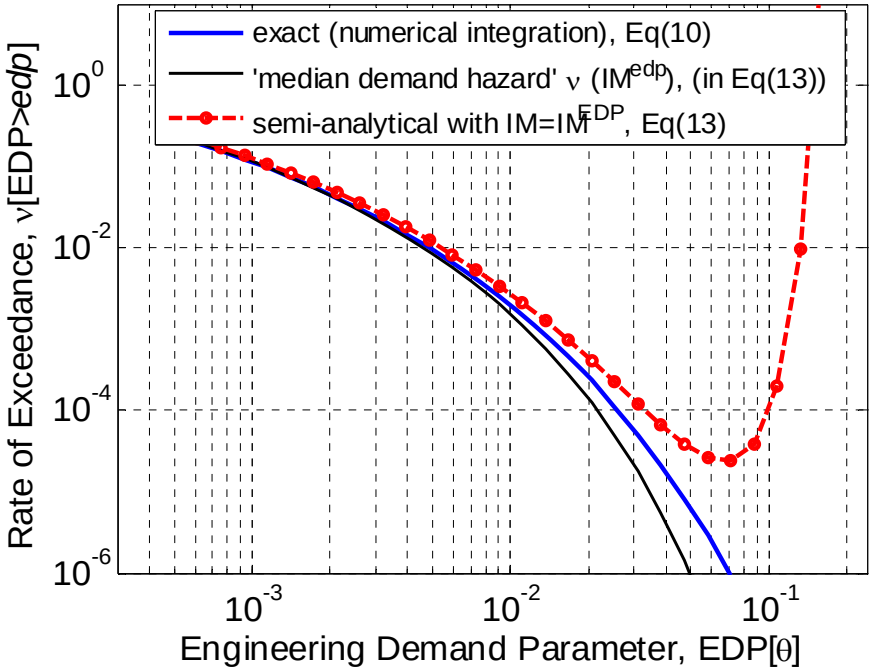


Figure 3-6: Semi-analytic closed-form solution using $IM = IM_{edp}$.

This problem of what slope, k_{eq} (or k), to use to compute $\nu(edp)$ is not just a problem with using the hyperbolic model (Equation (3-4)), but is also a problem when using the closed-form solution for the power law model (Equation (3-11)). This problem can be resolved if the closed form solution is inverted, therefore becoming a function of ν. Hence, the solution procedure is as follows: (i) for the selected rate of exceedance calculate the log-log slope, k_{eq}, of the hazard curve; and (ii) calculate the *EDP* using the selected rate of exceedance and k_{eq}. Since both the approximation for the slope and the corresponding *EDP* are calculated based on the same rate of exceedance, then the solution should be more accurate, without requiring iteration.

Inverting the closed form solution proposed by Luco and Cornell [22] (Equation (3-9)) gives the *EDP* for a given rate of exceedance, ν:

$$EDP(\nu) = a\left(\frac{\nu}{k_0}\right)^{-\frac{b}{k}} \exp\left(\frac{1}{2}\frac{k}{b}\beta^2_{EDP|IM}\right) \tag{3-13}$$

Calculation of *EDP*(ν) using the hyperbolic hazard model can be determined by starting with Equation (3-11), where $\nu\left(IM^{edp}\right)$ is given by substituting Equation (3-10) into Equation (3-4). Then the log-log slope of the hazard curve, k, can be expressed as a function of ν by substituting the second term of Equation (3-4) into Equation (3-12). Finally, this modified form of Equation (3-11) can be rearranged as a function of *EDP* to yield:

$$EDP(\nu) = aIM_{asy}^{b} \exp\left[\frac{\alpha b}{V - \frac{V^4}{2b^2\alpha^2}\beta_{EDP|IM}^2}\right] \tag{3-14}$$

where $V = ln(\nu/\nu_{asy})$, and all other parameters have their usual meaning.

Figure 3-7 illustrates the approximation of Equation (3-14) to the 'exact' solution obtained via numerical integration, compared with the closed form solution using the power law hazard model (Equation (3-13)). It can be seen that the semi-analytical closed form solution displays far superior accuracy over a larger range of *EDP*, in comparison to the power law closed form solution. The semi-analytical solution is not exactly the same as the numerical solution however. The reason for this is that the convolution integral (e.g. Equation (3-8)) is dependent on the hazard at various intensity levels (i.e. the integration is over the full range of *IM*) and therefore at these different intensity levels the value of *k* varies. The semi-analytical solution, however, uses only a constant *k* value corresponding to the rate considered. For the bridge pier case study drift hazard curves (Figure 3-7) this over prediction ranged from 5-10% with a mean value of 7%. Even despite this inconsistency, the accuracy is sufficient considering the simplification of a closed form solution over numerical integration.

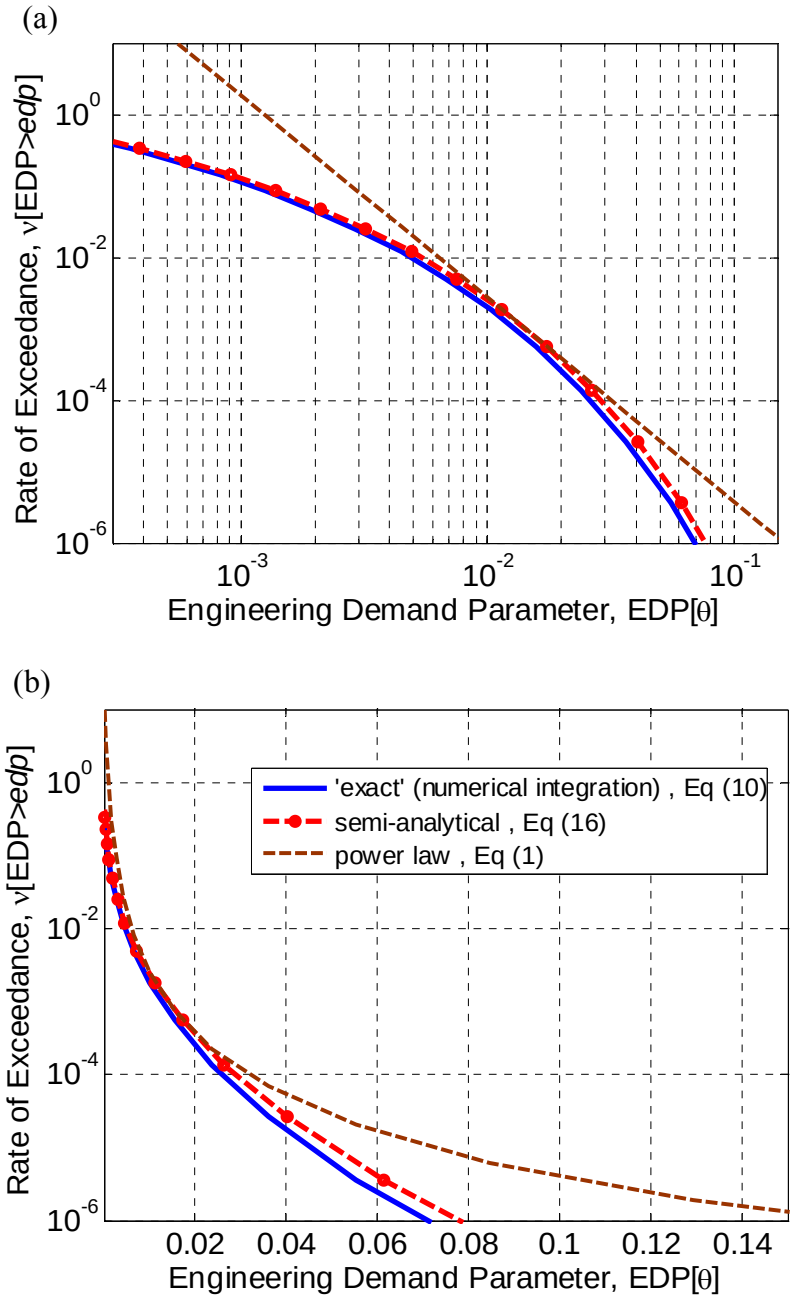


Figure: 3-7: Semi-analytic closed-form solution: (a) in ln*(IM)*–ln*(ν)* space; and (b) in IM–ln*(ν)* space.

3.7 Discussion

The previous sections have illustrated that probabilistic seismic demand analysis (PSDA) carried out analytically using the proposed hyperbolic model gives results which are significantly more accurate compared to those obtained using the closed form solution presented by others [4-8].

Note that the procedure to obtain the semi-analytical solution presented is not unique to

the seismic hazard model used. For example, the authors investigated the use of more classical probability distributions to model the hazard, such as the lognormal cumulative distribution function (CDF), i.e:

$$\nu = 1 - \left[\frac{1}{2} + \frac{1}{2} erf\left(\frac{\ln IM - \mu}{\sqrt{2}\sigma} \right) \right] \tag{3-15}$$

where *erf* = is the error function; and μ, σ are the logarithmic mean and standard deviation of the lognormal CDF. The same procedure can be used to obtain the solution for *EDP*(ν). The final form of *EDP*(ν) is significantly more complicated, and is given by:

$$EDP(\nu) = a \exp\left[b\mu + \sqrt{2} b \sigma erfinv\left(1 - 2\nu \exp\left\{ -\frac{1}{2} \frac{k^2}{b^2} \beta^2 \right\} \right) \right] \tag{3-16}$$

$$k = \frac{1}{\sigma\nu\sqrt{2\pi}} \exp\left(- \{erfinv(1-2\nu)\}^2 \right)$$

where *erfinv* = the inverse error function. Also, the ability of the lognormal distribution in approximating the hazard curve is somewhat limited (in certain cases) compared to the proposed hyperbolic model, in that it only has two parameters (mean and standard deviation) compared to the three for the hyperbolic model, making its fit of the hazard data less flexible.

As it has been shown that the 'curvature' of the hazard model can now be incorporated into the closed form solution, the only limitation of Equation (3-14) (or Equation (3-16) for that matter), is that the demand model parametric equation (Equation (3-7)) does not typically model the response in the region of global collapse well [6]. Shome and Cornell [25] showed that by using a power law function to describe the probability of collapse, a closed form solution could be obtained which considers collapse in the demand model. The (simple) inclusion of collapse using the proposed hyperbolic model (or lognormal model) could be the subject of further work.

3.8 Conclusions

A novel seismic hazard model has been developed which is non-linear in log-log space. The model, based on a hyperbolic shape in lnν - ln*IM* space, is typically fitted to seismic hazard data via least squares regression, and allows for the incorporation of epistemic uncertainty. The applicability of the model to seismic hazard data in New Zealand was illustrated for both peak ground acceleration, *PGA*, and 1.5 second spectral acceleration,

Sa(1.5s), and results for *PGA* were tabulated.

Via a performance-based assessment of a bridge pier designed to New Zealand standards, it was illustrated that the power law model for the seismic hazard significantly over-estimates the demand hazard if used over a wide range of the engineering demand parameter, *EDP*, and the proposed hyperbolic seismic hazard model was shown to be a much more reliable option.

A key benefit of the power law hazard model is that by combining it with a power law model for the relationship between *EDP* and *IM*, a closed-form solution for the demand hazard can be obtained. While the mathematical form of the proposed hyperbolic model does not permit such a solution, a semi-analytical solution procedure to calculate the drift hazard in closed-form was proposed. The procedure, while not exact, had a mean relative over-prediction of 7% for the given example. The proposed procedure is therefore a viable alternative compared to the closed-form solution utilizing the power law hazard model.

3.9 Acknowledgements

The authors would like to acknowledge Dr Graeme McVerry, from the Institute of Geological and Nuclear Sciences (GNS) for providing the PSHA data for New Zealand regions.

3.10 References

[1] Esteva L. Criteria for the construction of spectra for seismic design. 3rd Panamerican Symposium on Structures, Caracas, Venezula, 1967.

[2] Esteva L. Bases para la formulaci´on de decisiones de dise˜no s´ısmico. Ph.D. Thesis, Universidad Nacional Aut´onoma de M´exico, 1968,

[3] Cornell CA. Engineering seismic risk analysis. Bulletin of the Seismological Society of America 1968; 58(5): 1583–1606.

[4] Sewell RT, Toro GR, and McGuire RK. Impact of ground motion characterisation on conservatism and variability in seismic risk estimates. U.S. Nuclear Regulatory Commission, Washington, DC, 1991.

[5] Kennedy RP and Short SA. Basis for Seismic Provisions of DOE-STD-1020. Lawrence Livermore National Laboratory and Brookhaven National Laboratory, Washington, D.C., 1994.

[6] Cornell CA. Reliability-based earthquake-resistant design—the future, in 11th World Conference on Earthquake Engineering, Acapulco, Mexico, 1996.

[7] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[8] Jalayer F. Direct probabilistic seismic analysis: Implementing non-linear dynamic assessments. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2003, 173 pp.

[9] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges. Engineering Structures 2007; 29(10): 2662-2672, DOI: 10.1016/j.engstrut.2006.12.015.

[10] Lee TH and Mosalam K. Probabilistic seismic evaluation of reinforced concrete structural components and systems. Stanford University, Stanford, CA, 2006. http://peer.berkeley.edu

[11] Kunnath SK, Larson L, and Miranda E. Modelling considerations in probabilistic performance-based seismic evaluation: case study of the I-880 viaduct. Earthquake Engineering and Structural Dynamics 2006; 35(1): 57–75.

[12] Kramer SL. Geotechnical Earthquake Engineering. Prentice-Hall: Upper Saddle River, NJ., 1996; 653.

[13] Kulkarni RB, Youngs RR, and Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. 8th World Conference on Earthquake Engineering,, San Francisco, CA, 1984.

[14] Bommer JJ, Scherbaum F, Bungum H, Cotton F, Sabetta F, and Abrahamson N. On the use of logic trees for ground-motion prediction equations in seismic hazard assessment. Bulletin of the Seismological Society of America 2005; 95(2): 377–389, DOI: 10.1785/0120040073.

[15] Stirling MW, McVerry GH, and Berryman KR. A new seismic hazard model for New Zealand. Bulletin of the Seismological Society of America 2002; 92(5): 1878–1903.

[16] Standards New Zealand. NZS 3101 1995: Part 1: Concrete Structures Standard. Wellington, NZ, 1995.

[17] Carr AJ. Ruaumoko: Inelastic Dynamic Computer Program. Department of Civil Engineering, University of Canterbury, Christchurch, New Zealand, 2004.

[18] Otani S. SAKE, a computer program for inelastic response of R/C frames to earthquakes. University of Illinois at Urbana-Champaign, 1974.

[19] Dhakal RP, Mander JB, and Mashiko N. Bidirectional pseudodynamic tests of bridge piers designed to different standards. Journal of Bridge Engineering 2007; 12(3): 284-295.

[20] Solberg KM. Experimental and analytical investigations into the application of damage avoidance design. M.E. Thesis, University of Canterbury, 2007,

[21] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[22] Luco N and Cornell CA. Seismic drift demands for two SMRF structures with brittle connections. Structural Engineering World Wide, 1998.

[23] Vamvatsikos D and Cornell CA. Applied incremental dynamic analysis. Earthquake Spectra 2004; 20(2): 523–553.

[24] Vamvatsikos D and Cornell CA. Incremental dynamic analysis. Earthquake Engineering and Structural Dynamics 2002; 31(3): 491–514.

[25] Shome N and Cornell CA. Structural seismic demand analysis: Consideration of "Collapse", in 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, University of Notre Dame, South Bend, Indiana, 2000, 7.

[26] Bradley BA, Dhakal RP, and Mander JB. Parametric structure specific seismic loss estimation. 4th International Conference on Urban Earthquake Engineering, Tokyo, Japan, 2007.

[27] Deierlein GG, Krawinkler H, and Cornell CA. A framework for performance-based earthquake engineering, in 7th Pacific Conference on Earthquake Engineering, Christchurch, New Zealand, 2003.

4.Error Estimation of Closed-form Solution for Annual Rate of Structural Collapse

Bradley BA, Dhakal RP. Error estimation of closed-form solution for annual rate of structural collapse. *Earthquake Engineering and Structural Dynamics* 2008; **37**(15): 1721-1737.

4.1 Abstract

With the increasing emphasis of performance-based earthquake engineering (PBEE) in the engineering community, several investigations have been presented outlining simplified approaches suitable for performance-based seismic design (PBSD). Central to most of these PBSD approaches is the use of closed-form analytical solutions to the probabilistic integral equations representing the rate of exceedance of key performance measures. Situations where such closed-form solutions are not appropriate primarily relate to the problem of extrapolation outside of the region in which parameters of the closed-form solution are fit. This study presents a critical review of the closed form solution for the annual rate of structural collapse. The closed form solution requires the assumptions of lognormality of the collapse fragility and power model form of the ground motion hazard, of which the latter is more significant regarding the error of the closed-form solution. Via a parametric study, the key variables contributing to the error between the closed-form solution and solution via numerical integration are illustrated. As these key variables can not be easily measured it casts doubt on the use of such closed-form solutions in future PBSD, especially considering the simple and efficient nature of using direct numerical integration to obtain the solution.

4.2 Introduction

Performance-based earthquake engineering (PBEE) and performance-based seismic design (PBSD) concepts are growing in popularity amongst the engineering community because of their consistent nature with respect to the socio-economic aims of seismic design. PBEE and PBSD also allow for incorporation of the uncertainties in all aspects of seismic design and assessment. The growing importance of PBEE and PBSD is illustrated by its inclusion in recent significant documents [1-5] .

Typical key performance measures in PBEE include the annual rate of exceedance of a given level of demand or financial loss, and the annual rate of structural collapse. The direct incorporation of uncertainties in the aforementioned performance measures results in an integral equation, which is an application of the total probability theorem. In such equations, a cumulative density function (CDF) is integrated over all intensities with the ground motion hazard curve for a specific site.

A key concept advocated by researchers in this area is that for PBSD to be accepted in design, simplified methods must be available which allow reasonably accurate evaluations to be made based on sound underlying assumptions. For the aforementioned key performance measures, numerous references are available for 'closed-form' analytical solutions (e.g. Appendix A). The first closed-form solutions were published for the demand hazard in References [6, 7], and using similar assumptions, annual frequencies of limit state exceedance and structural collapse can also be computed [8-10].

Such closed-form solutions have been used extensively since their development. Cornell *et al*. [9] used the closed-form drift hazard solution in a load and resistance factor design (LRFD) approach, which is implemented in FEMA-350 [3]. Mackie and Stojadinovic [11] used closed-form solutions for damage and loss limit states to propose a PBSD approach for bridges. Zareian and Krawinkler [10] used the closed form solution for the annual rate of collapse, to propose a PBSD methodology considering structural collapse. The above three implementations also separate epistemic and aleatory uncertainties in the structural response and use the mean ground motion hazard curve. These two treatments allow the determination of the mean annual rate of exceedance of a particular performance measure with a specified level of confidence.

The closed-form solution for the annual exceedance rate of demand (i.e. demand hazard) is based on the following three assumptions: (i) the ground motion hazard curve is

approximated by a linear line in log-log space; (ii) the median demand given intensity is a linear function in log-log space; and (iii) the demand given intensity distribution is assumed lognormal with constant logarithmic standard deviation (herein referred to as the 'dispersion') over the range of intensity that is of interest. Because these assumptions are made only in the regions of interest of the relationships, then the resulting closed-form solution may be considered as a 'local approximation' of the key performance measure around the region of interest. For example, it is stated in Kennedy and Short [6] that "over any ten-fold difference in exceedance probabilities, such hazard curves may be approximated by the PSDA analytical equation".

Aslani and Miranda [12] compared the closed-form solution for the demand hazard with that obtained by direct numerical integration using parametric relationships for the mean and dispersion of the demand given intensity relationships. They illustrated the resulting error in the demand hazard curves due to each of the three aforementioned assumptions required in the closed-form solution becomes significant as the demand levels become significantly different from those which the parameters were fit too.

Recently, Bradley *et al*. [13] proposed a 'hyperbolic' parametric equation to represent the ground motion hazard which is significantly more accurate over a larger range of exceedance frequencies than the power-model equation used to obtain the closed form solution for the demand hazard (i.e. [6, 7]). It was then illustrated how a semi-analytical solution for the demand hazard could be obtained using the 'hyperbolic' hazard model. This work offered a potential solution to the problem of 'extrapolation' of the local approximation of the closed-form demand hazard solution to a larger range of exceedance frequencies. However, in computing the exceedance rate of a single value of demand, the semi-analytical solution of Bradley *et al*. [13] and the closed-form solution using the power-model equation are identical. Also, the semi-analytical solution given by Bradley *et al*. [13] still requires the two assumptions for the demand given intensity relationship which also introduce some extrapolation error [12].

From the above discussions, it is clear that the criticism of the closed-form solutions is primarily due to their inability to accurately extrapolate outside the immediate range over which the parametric relationships are fit. This chapter investigates the error in the closed-form solution for the annual rate of structural collapse (collapse hazard), which does not suffer from the problems of extrapolation as the demand hazard mentioned above; implications related to the demand hazard are also briefly addressed. Deaggregation [14-16] of the integral equation is used to determine the regions of ground motion intensity which significantly

contribute to the numerical value of the collapse hazard. Via a parametric study, key features of the integral equations that contribute to the error between the closed-form and exact numerical solutions are identified. Various means of fitting the power-model equation to the ground motion hazard data are discussed in light of the resulting errors in the parametric study.

4.3 Closed-form solution for the annual rate of structural collapse

Firstly, the Pacific Earthquake Engineering Research (PEER) Centre PBEE framework terminology is adopted herein. Therefore, seismic demand is referred to as an engineering demand parameter (EDP), and ground motion intensity as an intensity measure (IM). The 5% damped elastic spectral acceleration at the fundamental period of the structure ($Sa(T_1,5\%)$, herein Sa for brevity) is used as the IM.

The annual rate of structural collapse (collapse hazard) can be obtained by integrating (over the entire range of ground motion intensity) the conditional probability of collapse for a given level of intensity with the incremental probability of occurrence of that ground motion intensity. The mathematical formulation of the collapse hazard is given in Equation (4-1), which is an application of the Total Probability Theorem [17]:

$$\lambda_C = \int_0^\infty P\left(C \middle| IM = im\right) \left| \frac{d\lambda(IM > im)}{dIM} \right| dIM \tag{4-1}$$

where λ_C = the annual rate of collapse; $P(C \mid IM = im)$ = the conditional probability of collapse given $IM = im$ (collapse fragility curve); and $\lambda(IM > im)$ = the annual rate of exceedance of $IM = im$ (ground motion hazard) at the site. The absolute value signs around the derivative of the ground motion hazard are used as its value is negative.

In order to obtain a closed-form solution of Equation (4-1), several simplifying assumptions are required. Firstly, the intensity at which collapse is observed to occur is assumed to be of the form given in Equation (4-2):

$$\ln\left(IM \mid C\right) = \ln(\eta_Z) + \ln(\varepsilon_{UZ}) + \ln(\varepsilon_{RZ}) \tag{4-2}$$

where η_Z = the median IM causing collapse; and $\ln\left(\varepsilon_{RZ}\right)$ and $\ln\left(\varepsilon_{UZ}\right)$ are aleatory and epistemic uncertainties having a normal distribution with zero mean and standard deviations of β_{RZ} and β_{UZ}, respectively. Equation (4-2) results in a collapse fragility curve (due to aleatory randomness) which has a cumulative lognormal distribution, and η_Z also having a lognormal distribution.

The ground motion hazard is also assumed to have a linear form in log-log space given by Equation (4-3):

$$\ln\left(\lambda(IM)\right) = \ln(k_0) - k.\ln(IM) + \ln(\varepsilon_{UIM}) \tag{4-3}$$

where k_0 and k are constants fitted to the ground motion hazard in the region of interest [13], and $\ln(\varepsilon_{UIM})$ is a normal random variable with zero mean and standard deviation β_{UIM}, representing epistemic uncertainty in the ground motion hazard. Hence, the mean of Equation (4-3) is $\overline{\lambda(IM)} = k_0 IM^{-k}$. One further assumption is that $\ln(\varepsilon_{UZ})$ and $\ln(\varepsilon_{UIM})$ are independent of each other, but within each random variable there is a perfect correlation at various levels of intensity (e.g. $\ln(\varepsilon_{UZ})$ is perfectly correlated to itself at various levels of intensity).

Based on the aforementioned assumptions, the evaluation of Equation (4-1) using integration by parts leads to the following closed-form solution for the mean collapse hazard (See Jalayer [18] for details on a similar process to obtain the demand hazard):

$$E[\lambda_C] = k_0 \eta_Z^{-k} \exp\left[\frac{1}{2} k^2 \left(\beta_{UZ}^2 + \beta_{RZ}^2\right)\right] \tag{4-4}$$

Furthermore, λ_C is a lognormal random variable with dispersion:

$$\sigma_{\ln \lambda_C} = \sqrt{\beta_{UIM}^2 + k^2 \beta_{UZ}^2} \tag{4-5}$$

Equation (4-4) indicates that the expected value of the annual rate of collapse can be obtained from the annual rate of exceedance of the median *IM* value causing collapse, η_Z, and then a multiplying factor (the exponential term) which represents the effect of uncertainty on the annual rate of structural collapse. This factor indicates that increasing the uncertainty in the collapse fragility curve and the log-log slope of the ground motion hazard curve, increases the expected rate of collapse. In particular, it is noted that while increasing the dispersion of the collapse fragility curves increases the probability of collapse at *IM* values lower than the median *IM* but reduces the probability of collapse at *IM* values larger than the median *IM*, it is the small IM values which occur significantly more frequently.

Figure 4-1a gives a typical probabilistic seismic demand analysis (PSDA) plot which has been derived via Incremental Dynamic Analysis (IDA) [19] of a single-degree-of-freedom (SDOF) model of a New Zealand bridge pier. The SDOF model uses a lumped plasticity (frame) element with the modified Takeda hysteresis having both strength and stiffness degradation. Further details on the bridge structure and its modelling can be found in Reference [20]. Each of the lines in Figure 4-1a represent the result of an individual record

scaled over a range of *IM*, and the dots at the end of the lines represent the projection (to the right boundary of the figure) of the last *IM* value before structural collapse was observed. Structural collapse is considered as the limit state of global sidesway instability (indicated numerically by non-convergence of the analysis). Global collapse associated with loss of vertical carrying capacity (due to axial and/or shear failures) is not considered here due to the lack of reliable analysis tools for capturing such phenomena [10]. Others have also defined global collapse when the slope of the tangent of the IDA curve drops below 20% of the initial tangent [3, 21], but this was not done here. Based on the sample mean and standard deviation of the *IM*'s causing collapse, a lognormal distribution of collapse given *IM*, can be defined, which is also shown in Figure 4-1a. Figure 4-1b gives a typical comparison between the seismic hazard curve for Wellington, New Zealand, and the approximation of the power-model (Equation (4-3)), fitted tangentially to the median *IM* causing collapse of the bridge structure considered.

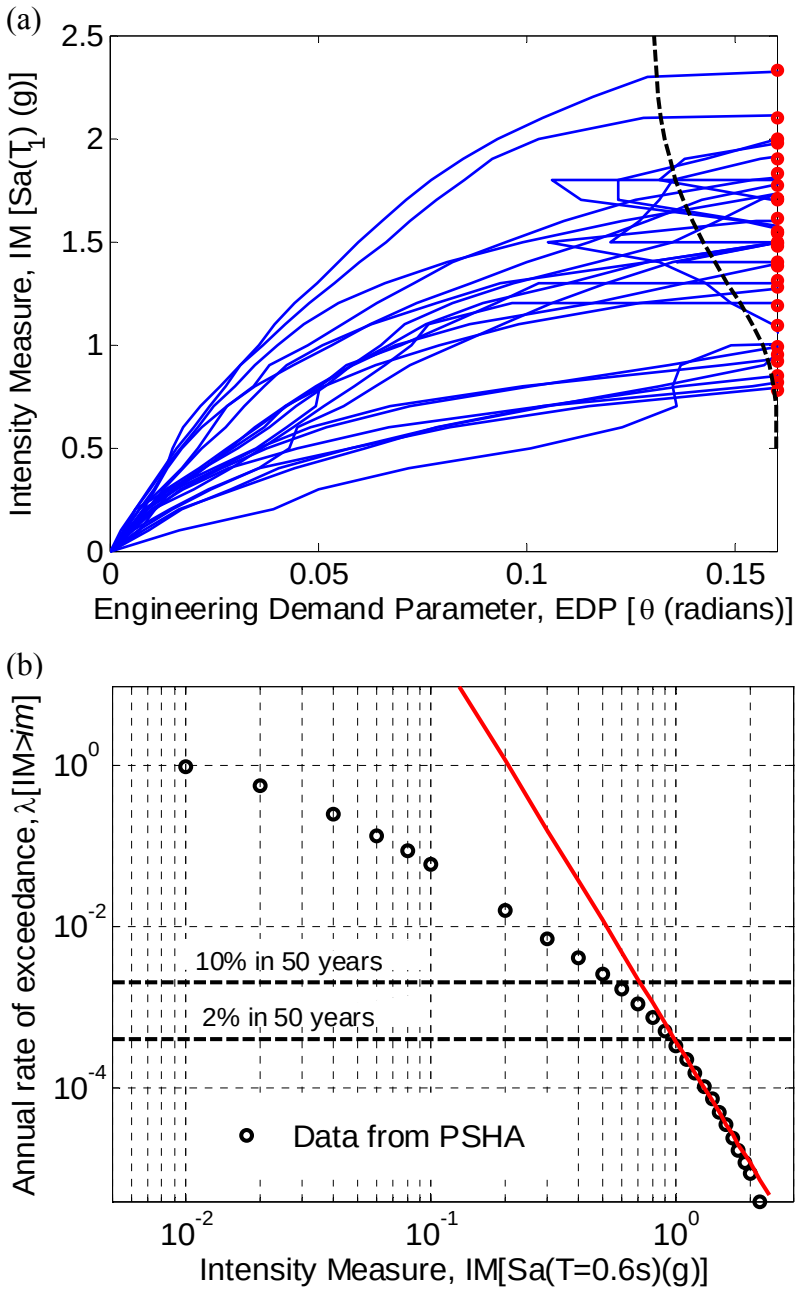


Figure 4-1: Comparisons between observational data and the parametric equations for the closed-form solution: (a) Seismic intensity-collapse relationship; and (b) Ground motion hazard.

4.4 Sources of error in collapse hazard closed-form solution

Firstly, discussions are restricted to the error associated with the expected value of the collapse hazard (i.e. Equation (4-4)), and consider only one source of uncertainty in the collapse fragility curve. This uncertainty may be solely aleatory, or a square-root-sum-of-squares (SRSS) [10] combination of both aleatory and epistemic uncertainties.

Equation (4-1) illustrates that the collapse hazard is a function of both the collapse

fragility curve and the derivative of the ground motion hazard curve. Figure 4-2 gives a comparison of the lognormal collapse fragility curve and the empirical CDF based on the IDA data in Figure 4-1a. It can be seen that the typical [10] lognormal approximation is acceptable for this data, based on Kolmogorov-Smirnov (K-S) goodness of fit test [17]. Various other studies have illustrated that this assumption is adequate and it has been used via direct numerical integration with the full representation of the seismic hazard of the site [10, 12]. A non-parametric form of the collapse fragility can be used, however care should be taken to ensure that enough ground motions are used such that the annual rate of structural collapse is not sensitive to the 'steps' in the empirical CDF. As an alternative to developing collapse fragility curves via IDA data, various data are available for collapse capacities for generic moment resisting frames and shear walls, which are useful for preliminary design assessments [22].

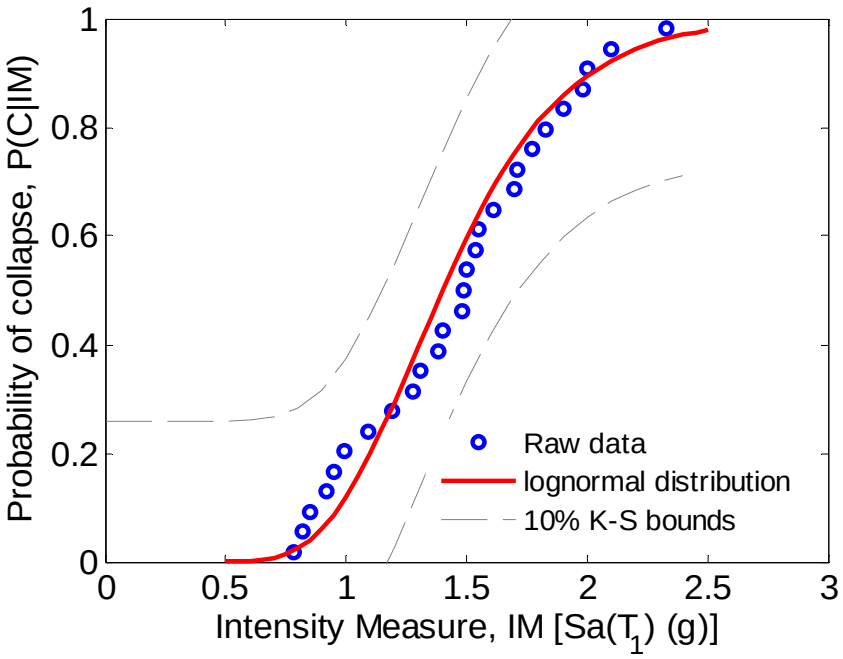


Figure 4-2: Collapse Fragility curve for the IDA curves in Figure 4-1a.

Based on the above discussion, as the lognormal assumption for the collapse fragility curve is adequate, it will be shown that the most restrictive assumption in order to derive Equation (4-4) is the power-model approximation of the ground motion hazard curve. The power-model therefore assumes that the ground motion hazard is linear in log-log space which is considered as a 'local approximation'. The potential error comes from the fact that as Equation (4-1) involves integration over the entire range of *IM*, the power-model solution will potentially inaccurately approximate the likelihood of ground motions of $IM = im$ occurring over a large range of *IM*. This potential inaccurate approximation is due to the typical

'concave from below' shape of ground motion hazard curves in log-log space [13], compared with the linear (in log-log space) curve of the power-model.

As the power-model assumes that the ground motion hazard is linear in log-log space, the error will likely be a function of the 'curvature' of the hazard curve. Here, 'curvature' (ϕ) is defined as the second derivative of the ground motion hazard curve in log-log space (i.e. the rate of change of the tangential slope, k). Because Equation (4-1) combines the ground motion hazard curve with the cumulative probability of collapse, the major contribution to the integral will occur from ground motion intensities around the central IM value causing structural collapse, η_Z. For example, in the limiting deterministic case (when there is no uncertainty), only η_Z is used to evaluate Equation (4-1). The range of IM values that significantly contribute to the integral (and hence the error in the closed-form solution) will therefore be a function of the likelihood of these IM values causing collapse to occur, $P(C \mid IM = im)$. Hence, any error in the closed-form solution (Equation (4-4)) will also be a function of the dispersion in the collapse fragility curve (herein denoted simply as β).

4.5 Parametric study on error in closed-form solution using a tangent-fit to hazard data

To investigate the effects of curvature, ϕ, and dispersion, β on the error in the closed-form solution, a parametric study was carried out which is described in the following paragraphs. For brevity, the term 'hazard' will be used in reference to 'ground motion hazard'. Note that both the closed-form solution and 'exact' numerical integration solution compared here use the lognormal assumption for the collapse fragility curve (i.e. not the raw data depicted in Figure 4-2). Therefore, differences between the outcomes of these two approaches are solely due to the representation of the ground motion hazard curve.

To obtain an estimate of the curvature of the hazard curve around the region of interest, the parametric form for the ground motion hazard model proposed by Bradley *et al*. [13] is used, which is given by:

$$E\left[\ln\left(\lambda(IM)\right)\right] = \ln\left(\lambda_{asy}\right) + \alpha\left[\ln\left(IM / IM_{asy}\right)\right]^{-1} \tag{4-6}$$

where λ_{asy}, IM_{asy}, and α are constants to be fit by nonlinear regression. For the above parametric form the curvature at a given point can be found from:

$$\phi = \frac{\partial^2[\ln(\lambda(IM))]}{\partial[\ln(IM)]^2} = 2\alpha\left[\ln\left(\frac{IM}{IM_{asy}}\right)\right]^{-3} = \frac{2}{\alpha^2}\left[\ln\left(\frac{\lambda(IM)}{\lambda_{asy}}\right)\right]^3 \quad (4\text{-}7)$$

where the central and right-hand side algebraic expressions are the curvature as a function of IM and λ, respectively. Herein, unless otherwise stated, ϕ is calculated at the median IM corresponding to the collapse, η_Z.

To account for the fact that this simple definition of curvature will not be an exact measure of the error, five hazard curves for the major centres in New Zealand [23] were used. The hazard curves for these five regions, along with their curvatures as a function of rate of exceedance are presented in Figure 4-3a and 4-3b, respectively. It can be seen that these hazard curves represent a wide range of site seismicity, from low in Auckland, to high in Otira. Figure 4-3b shows that the curvature of the hazard curves increases as the rate of exceedance reduces. It is also interesting to note that the curvature of the hazard curves is not directly related to the seismicity of the site. For example, the Christchurch hazard has a far larger curvature than the Wellington hazard, despite the Wellington hazard having a larger seismicity. A similar comparison between the Auckland and Dunedin hazards can also be made. As it will be shown later, the error in the closed-form solution increases as the curvature of the ground motion hazard increases. This indicates that the error is not directly related to the seismicity of the site.

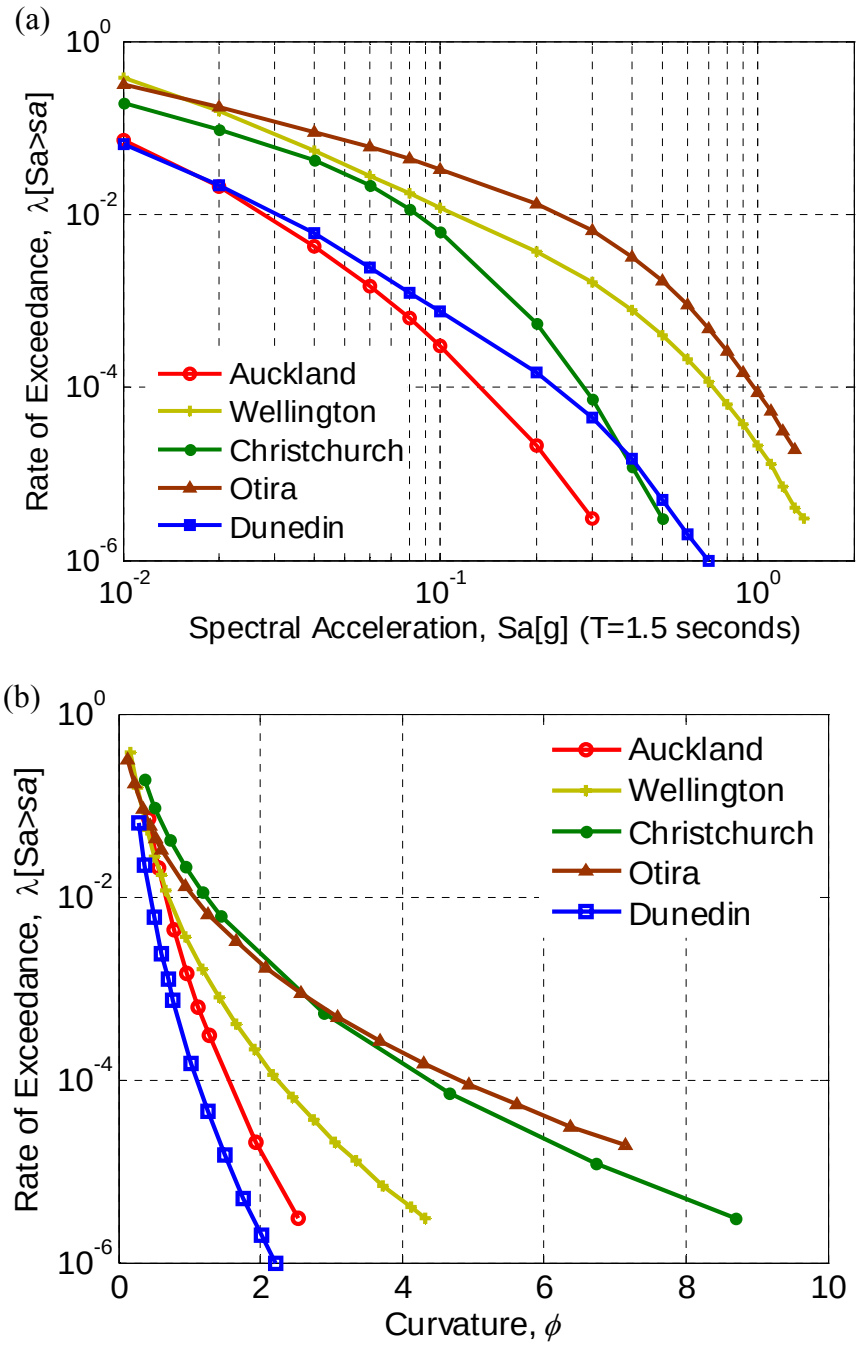


Figure 4-3: Hazard curves used in case study and curvature as a function of rate of exceedance: (a) Ground motion hazard curves; and (b) 'curvature' of hazard curves.

In order to illustrate that the error in the closed-form solution (Equation (4-4)) is a function of both ϕ and β the concept of deaggregation [14-16] is used. Deaggregation allows the contribution of different values of the integrand to the integral to be graphically illustrated. Figure 4-4 shows four deaggregation plots of Equation (4-1) using both the 'exact' numerical solution and the closed-form solution, where the parameters of the power-model of the ground motion hazard (Equation (4-3)) have been obtained by fitting the model tangentially to the

raw hazard data at $IM = \eta_Z$. In these figures, the Christchurch ground motion hazard curve (which is of moderate seismicity) was used. Two frequencies for the median *IM* causing collapse and two values of the dispersion were considered. The frequencies for the median *IM* causing collapse considered were $\lambda = 2.1\text{x}10^{-3}$ and $\lambda = 1\text{x}10^{-4}$ (i.e. from the ground motion hazard curve, the median *IM* causing collapse, η_Z, has these exceedance frequencies). These two frequencies represent the upper and lower ranges of likely collapse frequencies. For example, non-ductile flexure-shear critical structures typically have an annual rate of collapse which can be greater than $2.1\text{x}10^{-3}$ (e.g. [24]), while for current code-conforming structures the collapse hazard is typically lower than $1\text{x}10^{-4}$ (e.g. [25]).

The first dispersion value used was $\beta = 0.3$. This dispersion value would typically occur for 'efficient' [8, 26] *IM* such as the inelastic spectral displacement, S_{di}, proposed by Tothong and Luco [27]. The second value of $\beta = 0.5$ was used as a value representative of dispersions due to a relatively inefficient *IM* (such as elastic spectral acceleration, S_a, which is the most commonly used *IM*). For example, although not explicitly mentioned, the dispersion (due to aleatory uncertainty) in the collapse fragility (using the first mode spectral acceleration as the *IM*) given in Reference [10] is approximately 0.42. Other cases where a large dispersion may be measured could be where: (i) several designs are to be compared, which do not have the same characteristics (e.g. fundamental period), in which case the use of a structure-dependent *IM*'s (such as S_a) may not be appropriate (e.g. Reference [11] gives 13 dispersion values ranging from 0.33 to 0.56 for simple bridge structures using *IM* = *PGV*); (ii) higher-mode effects are important (e.g. in flexible structures an *IM* such as S_a may not accurately predict a multi-mode dominated response [28]); and (iii) near-fault velocity-pulse effects [28].

It can be seen in Figure 4-4 that as the curvature and dispersion increase so does the error between the closed-form solution and the 'exact' solution using numerical integration. Here, the error has been represented in the form of an error ratio, defined as:

$$E_{ratio} = \frac{\lambda_{C,approx}}{\lambda_{C,exact}} \tag{4-8}$$

where $\lambda_{C,approx}$ = the closed-form solution (Equation (4-4)); and $\lambda_{C,exact}$ = the 'exact' numerical solution of Equation (4-1). Figure 4-4a-Figure 4-4c therefore have errors of 16%, 60%, and 77%, respectively, while Figure 4-4d has a 7-fold (700%) error. It is also observed that the integration error contributed by *IM* values larger than the median *IM* causing collapse, η_Z, is negligible compared to the error contributed by *IM* values below η_Z. This consistent nature of the error in the closed-form solution potentially allows other means of fitting the

ground motion power-model which is discussed in the following section.

Based on typical values for the dispersion observed in the literature [7-13, 18-20, 26-30] and exceedance rates of collapse that could occur for a wide range of structures [24, 25], a parametric study was performed using $\beta = 0.2\text{-}0.6$ and $\lambda = 10^{-2}\text{-}10^{-5}$. The results of the parametric study are presented graphically in Figure 4-5. Figure 4-5a shows the error ratios (as defined in Equation (4-8)) for $\beta = 0.2$ and 0.3. The dashed lines surrounding the data points are used to clearly define the data points for each β value. The relatively small scatter between the data points for the five different hazard curves indicates that β and ϕ capture the salient features of the error between the closed-form solution and the 'exact' numerical solution. Figure 4-5b shows the results for $\beta = 0.4$ - 0.6. Again, the dashed lines are used to distinguish between different β values. It is obvious from both figures that the variation in error between the results for different hazard curves increases as β increases.

To give a practical viewpoint of Figure 4-5, consider the use of the closed-form solution with $\beta = 0.42$. This value of β is that (approximately) obtained in Reference [10], and is below the median of the β values used in Reference [11]. Assume that the structure is designed to current ductile design philosophy and has a fundamental period of T=1.5s (i.e. so that the hazard curves of Figure 4-4a are used), and a median collapse intensity, $\eta_Z = 1.4$g (this is slightly less than $\eta_Z \sim 1.75$g used in Reference [10], in which the structure had a period of T = 1.2s). Based on the results of Figure 4-5b the error ratios for the Wellington and Otira sites would be approximately 3.1 and 10.1, respectively. This means that if the collapse rate for Otira was found (using the closed-form solution) to be on the order of $\lambda = 10^{-4}$, then its actual value is likely to be in the region of $\lambda = 10^{-5}$. Note also, that the value of $\beta = 0.42$ represents aleatory uncertainty only. If epistemic uncertainty (which is typically in the region of 0.4-0.45 [10, 30] is also included in an SRSS form, then $\beta \sim 0.6$ and the error ratio will be in excess of 20. Such large errors defeat the purpose of using a probabilistic-based measure of performance.

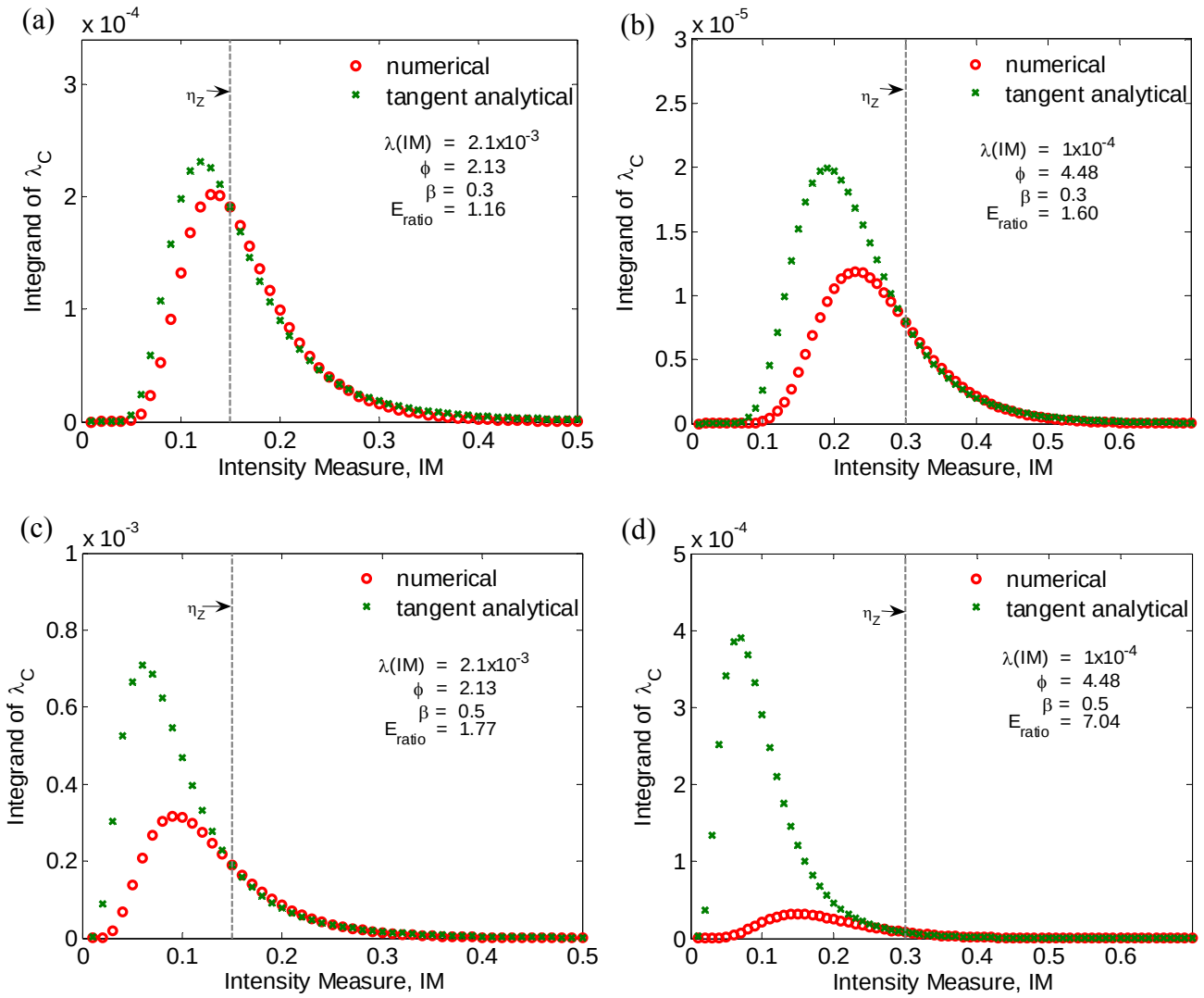


Figure 4-4: Deaggregation of Equation (4-1) for different values of φ and β

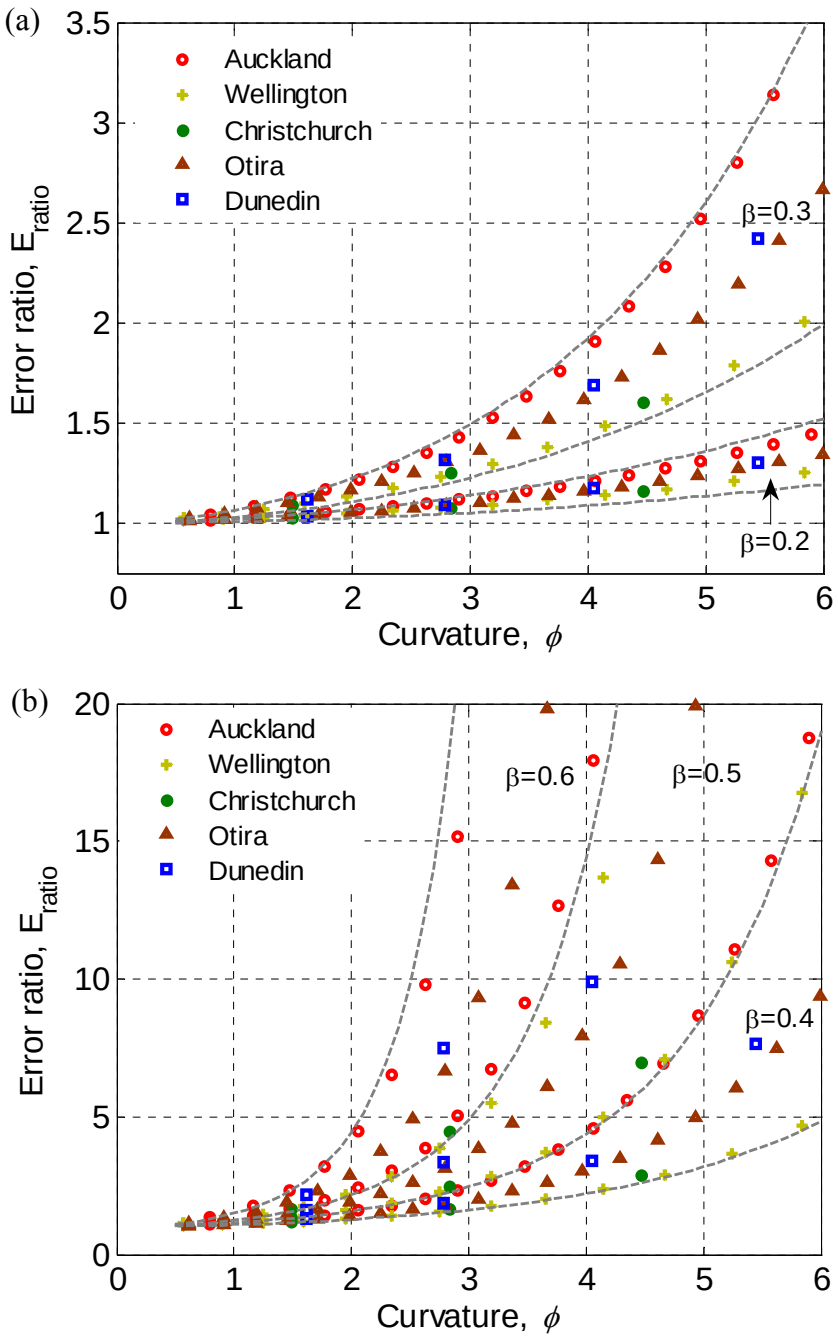


Figure 4-5: Error estimates for the tangent based approximation to the ground motion hazard curve for: (a) $\beta = 0.2$ and 0.3; and (b) $\beta = 0.4$,0.5, and 0.6.

4.6 Alternative non-tangent power-model fits to ground motion hazard

The deaggregation results of Figure 4-5 illustrated that using a tangent based fit of the hazard curve to determine k results in significant over-approximation of the contribution of ground motions with $IM < \eta_Z$. This occurs because the log-log slope of the hazard curve, k, is

too large over the region $IM < \eta_Z$. Therefore, a reduction in the value of k will likely reduce such an over-approximation. Such non-tangent methods have been suggested previously by others. For example, when computing the demand hazard around the design basis earthquake (DBE) and maximum considered earthquake (MCE) rate region, Jalayer [18] suggested fitting the power-model hazard as a secant through the DBE and MCE points of the ground motion hazard.

In this work, several alternative methods of fitting k were investigated, which include some of the following: (i) multiply the tangent-based fit of k by some constant; (ii) fit k tangential to the hazard curve at some rate less than η_Z; (iii) use a secant-based fit of k between two points either side of η_Z; and (iv) use regression over some region of the ground motion hazard to determine the power-model parameters. Table 4-1 gives a summary of the resulting error ratios for a selection of the different fitting methods used, for the Christchurch hazard. For example, using the secant-based fit with one point at $IM = \eta_Z$, and the other at a value of IM which has rate of exceedance equal to ten times that of η_Z (first row for base case (iii) in Table 4-1) resulted in relatively accurate (compared to the tangent-fit) results over the wide range of values (and different hazard curves) used in the parametric study. Figure 4-6a and 4-6b give the deaggregation plots obtained using the secant-based fitting of k at $IM = \eta_Z$ and $IM_{10\lambda}$, which are for the same (φ, β) scenarios as Figure 4-4b and 4-4d which used the tangential fit of k. The two vertical dashed lines in Figure 4-6a and 4-6b show the IM values through which the secant-fit was performed. In particular, for $\beta = 0.3$ and $\lambda = 10^{-4}$ (Figure 4-6a), the error ratio for the secant-based fit is 2% (Figure 4-6a) compared to the 77% error using the tangent-based fit (Figure 4-4b). From the discrepancies between the numerical and closed-form solutions relative to the points where the secant-fit was performed, it becomes obvious that for this type of fitting, the closed-form solution under-predicts the contribution from ground motion intensities with $IM > IM_{10\lambda}$ and over-predicts the contribution of ground motion intensities with (approximately) $IM < IM_{10\lambda}$. Hence, the accuracy reflected in the error ratio of 1.02 is the result of 'subtractive cancellation', that is errors in one region are negated by errors (of opposite nature) in another region. Obviously, over a large range of β and φ values it is unlikely that such 'subtractive cancellation' will consistently occur. This is illustrated in Figure 4-6b, where for $\beta = 0.5$, $\lambda = 10^{-4}$, the error ratio is 1.86; still a significant reduction however compared to the 7-fold error using the tangent-based solution (Figure 4-4d).

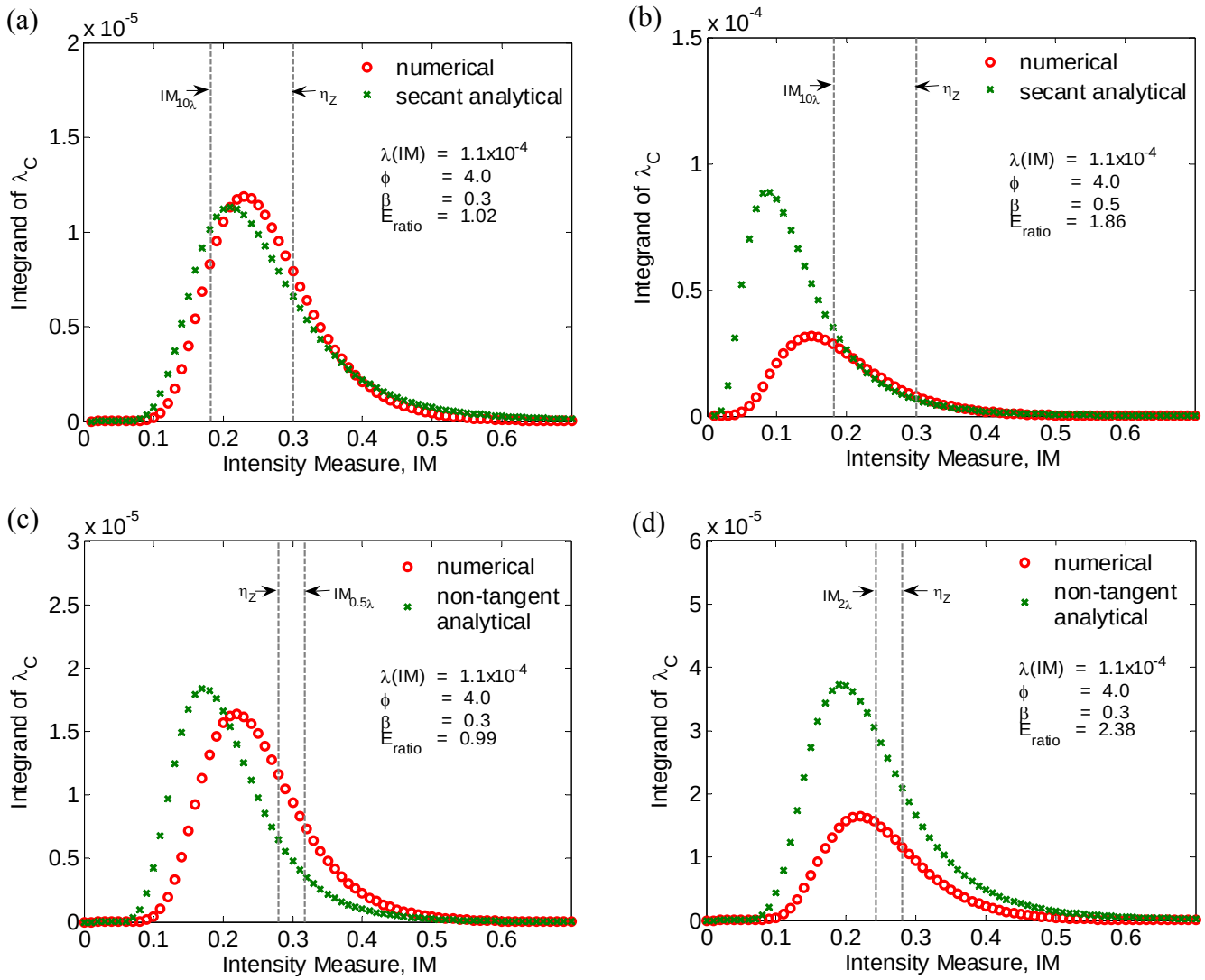


Figure 4-6: Deaggregation of Integral using secant-based power-model fit: (a) low error ratio due to subtractive cancellation; (b) non-occurrence of subtractive cancellation when β is increased; (c) low error ratio for $IM_{0.5\lambda}$ fit method; and (d) large error for $IM_{2\lambda}$ fit method.

Several other fitting methods such as 'k=0.75k_t' and '$IM_{0.5\lambda}$' from Table 4-1 appear to be more accurate, particularly at large values of β and φ. This however results from the aforementioned 'subtractive cancellation', and these results significantly under-predict the exact value for small β and φ, yielding error ratios of 0.93 and 0.67 for the (λ, φ, β) = ($2.1\text{x}10^{-3}$, 2.0, 0.3) scenario. It is also interesting to note that based on discussions in the previous section regarding the majority of the error ratio being contributed by $IM < \eta_Z$ one would expect that if the power-model hazard is fit tangentially at a rate greater than that of $IM = \eta_Z$ the error would be smaller than that which occurs when the power-model hazard is fit at a rate less than $IM = \eta_Z$. Thus, it would be expected that the $IM_{2\lambda}$ fitting method is better than the $IM_{0.5\lambda}$ fitting method. Rows 6 and 7 of Table 4-1 illustrates that this assumption is not correct,

in fact one would argue that based on Table 4-1 the $IM_{0.5\lambda}$ fitting method is better than the $IM_{2\lambda}$ fitting method. Figure 4-6c and 4-6d show the deaggregation of the collapse hazard for the case of (λ, φ, β) = ($1x10^{-4}$, 2.0, 0.3). Figure 4-64-6c illustrates that using the low error ratio for the $IM_{0.5\lambda}$ fitting method is due to 'subtractive cancellation' as the analytical solution under-predicts the contribution around the region where the power-model is fit, and over-predicts the region where $IM >> \eta_Z$. Figure 4-64-6d illustrates that in this case fitting the power-model tangentially at a rate less than that of the $\lambda(\eta_Z)$ results in over approximation of the integral over the entire range of IM values.

Table 4-1: Error ratios for various ground motion hazard curve fitting methods.

Type	Fit factor /location	$\varphi = 2.0$[1]		$\varphi = 3.0$[1]		$\varphi = 4.0$[1]	
		$\beta = 0.3$	$\beta = 0.5$	$\beta = 0.3$	$\beta = 0.5$	$\beta = 0.3$	$\beta = 0.5$
Base-case	tangent	1.16	1.78	1.28	2.72	1.51	5.49
(i)	k=0.85kt	1.01	1.20	1.03	1.47	1.08	2.14
	k=0.75k_t	0.93	0.96	0.92	1.03	0.88	1.24
(ii)	$IM_{0.5\lambda}$	0.67	1.34	0.78	2.39	0.99	5.95
	$IM_{2\lambda}$	2.05	2.54	2.15	3.39	2.38	5.67
(iii)	$IM_{10\lambda}, \eta_Z$	0.95	1.03	0.967	1.25	1.02	1.86
	$IM_{5\lambda}, IM_{0.2\lambda}$	0.89	1.31	0.99	2.02	1.20	4.09
(iv)	$\gamma = 1$	0.52	0.71	0.65	0.73	0.47	0.46
	$\gamma = 2$	0.79	1.01	0.67	1.24	1.65	1.00
	$\gamma = 2.5$	0.90	1.48	1.15	2.32	0.84	1.67
	$\gamma = 3$	0.98	1.65	1.27	3.11	1.02	2.69
	$\gamma = 3.5$	1.02	1.74	1.37	3.90	1.17	3.81

[1] φ = 2.0, 3.0 and 4.0 correspond to the Christchurch ground motion hazard at approximately $\lambda(IM) = 2.1x10^{-3}$, $3.5x10^{-4}$ and $1.0x10^{-4}$, respectively.

For the regression fitting method we solve the least squares optimisation problem with various weighting functions:

$$\text{Minimise } R = \sum_{i=1}^{n} w_i \left[\ln(\lambda_i) - \ln(\lambda(IM_i))\right]^2 \tag{4-9}$$

where λ_i = data points of ground motion hazard curve; $\lambda(IM_i)$ = value of λ obtained from the power-model parametric equation (Equation (4-3)); and w_i = the weighting function for data point i. It would seem logical that the weights would be directly proportional to the range of IM values which contribute to the integrand. This will be a function of the distance between the data point (IM_i, λ_i) and $IM = \eta_Z$, as well as the aleatory uncertainty in the collapse fragility curve, β. The weight will therefore be related to the number of standard deviations of IM points from η_Z. As the weight should reduce as the number of standard deviations increases

then we use the inverse of the number of standard deviations for the weighting function:

$$w_i = \left(\frac{\beta_{RZ}}{\ln(IM_i) - \ln(\eta_Z)} \right)^{\gamma} \tag{4-10}$$

where γ is a parameter which controls the degradation of the weights as the number of standard deviations increases which is varied in the analysis to follow. The value γ = 0 would give a uniform weight to all data points. It is found that values of γ from 1-3.5 produce reasonable approximations to the integral. Figure 4-7a illustrates the hazard curves which are obtained for several different γ values by determining the parameters of Equation (4-3) via the solution of Equation (4-9). It can be seen that as the value of γ increases the power-model hazard curve approaches the tangent to the raw ground motion hazard data. Table 4-1 (base case (iv)) gives the error ratios when these parameters for the power-model are used. The tabulated values are also shown graphically in Figure 4-74-7b. It is evident that as before the error ratios generally increase as a function of dispersion, β, however, the error ratio is no longer directly proportional to the curvature which occurred in the tangent fit case (this is also true for several of the other non-tangent fits in Table 4-1) This is due to the method employed to compute the curvature (which uses only the second derivative of the hazard in log-log space at a single point), which was adequate when using a tangent-based fit, but does not appear adequate here. From Figure 4-74-7b it is also seen that no clear value of γ gives error ratios consistently close to 1.0, although out of all of the values of γ, one would probably suggest that γ = 2.0 yields the best results.

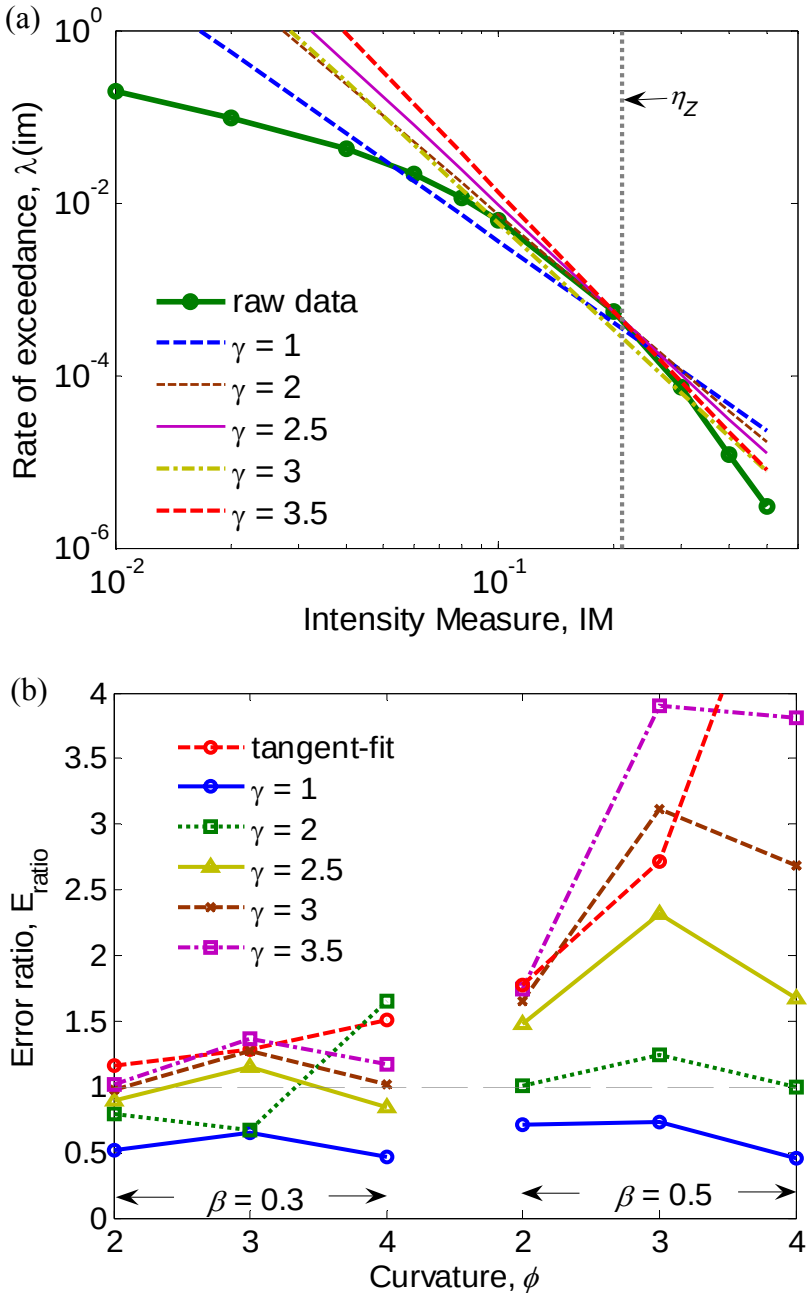


Figure 4-7: Illustration of using regression with various exponents for the weighting function: (a) effect on parameters, k, k_0 for $\beta = 0.5$ and $\varphi = 3.0$; and (b) error ratios.

4.7 Epistemic uncertainty in collapse hazard

As previously mentioned, when epistemic uncertainties are considered in (either or both of) the collapse fragility curve and the ground motion hazard curve, it is possible to compute the epistemic uncertainty in the collapse hazard, $\sigma_{\ln \lambda_c}$ (Equation (4-5)). Epistemic uncertainties arise in the collapse fragility due to finite sample uncertainty (estimating the parameters of the collapse fragility curve based on a finite number of points) and from

analysis modelling uncertainty (assumptions on soil-structure-interaction, hysteresis models, 3-dimensional effects etc.), while epistemic uncertainty in the ground motion hazard is due to assumptions in Probabilistic Seismic Hazard Analysis (PSHA) (e.g. type and parameters for attenuation relations, magnitude recurrence relationships etc.). In the following paragraphs examples are given of the computation of the epistemic uncertainty (and the resulting distribution) in the collapse hazard using both the exact and closed-form solutions.

To compute the epistemic uncertainty in the 'exact' numerical solution, 5000 Monte-Carlo (MC) simulations were used (which was checked manually to verify it was sufficient for convergence of the non-parametric distribution). In the MC simulation the median *IM* causing collapse, η_z, and the ground motion hazard, $\lambda(IM)$, are assumed to be lognormal random variables as stated to obtain the closed-form solution for $\sigma_{\ln \lambda_C}$ (Equation (4-5)). Figure 4-8a illustrates the empirical CDF using epistemic uncertainties of (β_{UZ}, β_{UIM}) = (0.4, 0.3) which are typical epistemic uncertainties appearing in literature [8, 34]. As the actual ground motion hazard is used in the exact solution (as opposed to the power-model approximation) the distribution of the collapse hazard no longer has a lognormal distribution (which is the case for the closed-form solution). It is seen in Figure 4-84-8a that while a lognormal distribution (based on the sample median and standard deviation) is an adequate approximation over the central region of the distribution, its accuracy diminishes toward the tails of the distribution. It is also apparent that the magnitude of the epistemic dispersion, $\sigma_{\ln \lambda_C}$, is significant (a value of $\sigma_{\ln \lambda_C}$ =1.75 means that assuming a lognormal distribution, the 84th percentile collapse rate is 33 times the 16th percentile collapse rate, and that the 90th percentile is 3.85 times more than of the mean). This large epistemic dispersion is consistent with the closed-form solution, in which the k^2 term amplifies the effect of the epistemic uncertainty in the collapse fragility curve, β_{UZ}.

It would seem intuitive that if the error ratio (E_{ratio}) in the closed-form solution for the expectation of λ_C is significant, then the error in $\sigma_{\ln \lambda_C}$ will also be significant. Of more importance however is: if the error λ_C from the closed-form solution is small, then will the error in $\sigma_{\ln \lambda_C}$ also be small? Possible reasons for significant error in $\sigma_{\ln \lambda_C}$ when E_{ratio} is small could be due to the aforementioned 'subtractive cancellation' in the expectation of the collapse hazard. Consider a single case using regression to fit the power-model (to the Christchurch hazard) with $\gamma = 2.0$, and using fragility and hazard parameters of $\beta_{RZ} = 0.3$, $\beta_{UZ} = 0.4$ and $\beta_{UIM} = 0.3$, $\varphi = 4.0$, respectively. These values are those used to obtain Figure 4-84-8a and from Table 4-1 give an error ratio of 1.0 for the expectation of the mean collapse

rate. Using the regression approach with $\gamma = 2.0$ gives $k = 3.79$, and thus Equation (4-5) gives $\sigma_{\ln \lambda_C} = 1.54$. This is a 12% error compared to the actual value of 1.75 given in Figure 4-84-8a. Figure 4-84-8b illustrates the effect of the underestimation of the dispersion on the distribution of the collapse hazard. It is evident that the error in the dispersion primarily induces error in the collapse hazard for smaller levels of confidence. For example, the 12% error in the dispersion (Figure 4-84-8b) gives an error of 150% in predicting the median (with respect to epistemic uncertainties) value of the collapse hazard.

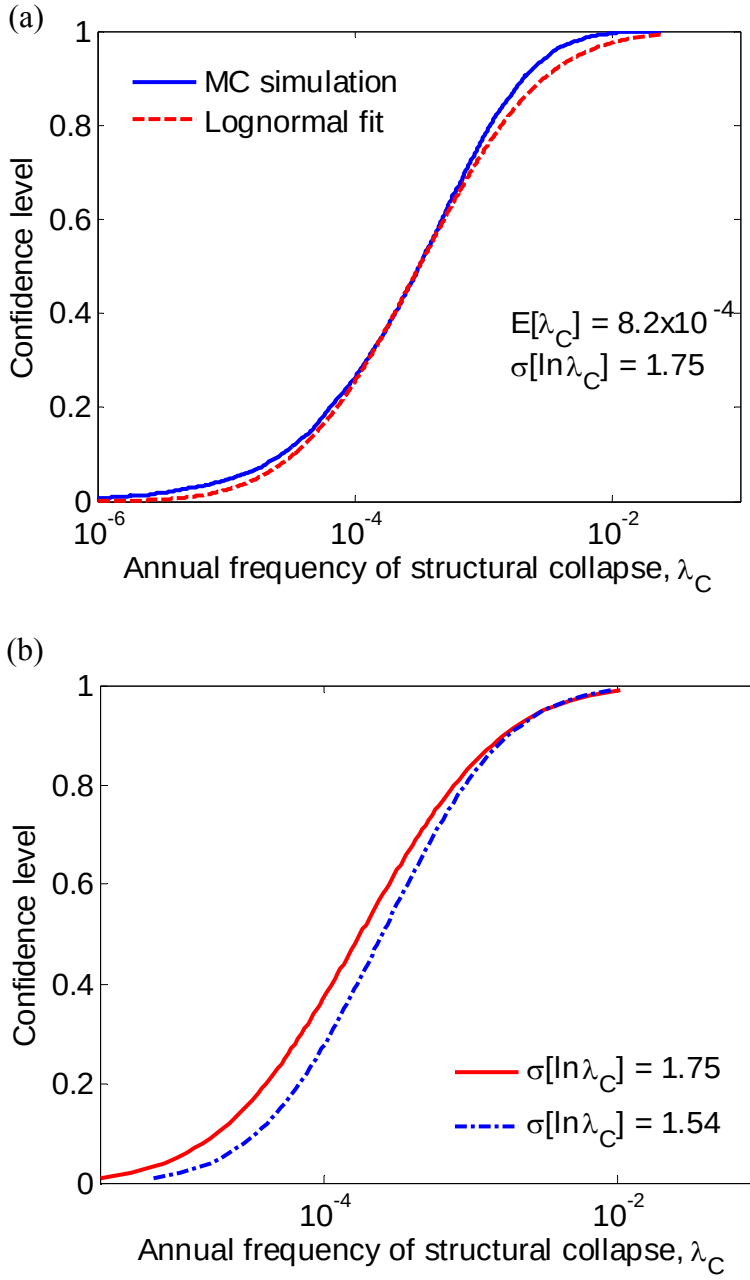


Figure 4-8: Error in collapse hazard distribution due to: (a) lognormal parametric assumption of non-parametric distribution; and (b) underestimation of the epistemic dispersion.

4.8 Discussion

Numerous methods have been considered for the determination of the parameters of the power-model ground motion hazard. Although for each specific scenario it is possible to find a method for determining the parameters which gives a small error ratio, it has been rigorously shown that no method in general is adequate over the large range of likely values of the factors primarily influencing the error. From these results it is apparent that the accuracy of the collapse hazard closed-form solution is very sensitive to the values of k and k_0 used (especially when the values of β and ϕ are significantly large). Hence if the closed-form solution is to be used then a great amount of care should be taken in selecting the values of these parameters. Based on the results of the parametric study it should be noted that there is unlikely to be any significant error when the closed-form solution is used to predict the annual rate of collapse for collapse-prone structures (i.e. those with an annual rate of collapse around $\lambda = 1\text{x}10^{-2}$). This is because Figure 4-3b illustrates that for frequent events, ϕ is typically less than 2, and Figure 4-5 shows that the error for this range of φ is small. Also, the error is strongly a function of the dispersion in the collapse fragility curve. This dependence on the dispersion further illustrates the need for advanced IM's which can accurately predict the effects of inelasticity and higher modes in complex structural behaviour [8, 26-28].

Another potential problem with the closed-form solution in its current form, as given in References [6-12], is that since the error is sensitive to the value of k used, in design environments either: (i) a large number of k values would have to be provided at different exceedance rates; or (ii) the raw hazard data would have to be provided, and designers should perform the necessary curve-fitting to obtain the value of k. It is likely, however, that the effort of the user to perform the power-model fit of the ground motion hazard (particularly if regression is used) is more than that required to directly numerically integrate Equation (4-1).

As a final remark, the results presented in this chapter for the error between the closed-form solution for the annual rate of collapse and the direct numerical solution are also insightful toward the errors in the closed-form demand hazard solutions given in [7, 12, 29]. It is already acknowledged from previous work (e.g. [12]) that the simplifying assumptions necessary for arriving at the closed-form solution of the demand hazard could lead to significant error if the region in which the local approximations are made is distant from the region of major contribution to the integral (i.e. extrapolation from the region of parameter fitting). This work has investigated the error in the closed-form solution for the collapse

hazard where the above comments regarding extrapolation are not applicable. It has been shown that no method of determining the parameters of the power-model for the ground motion hazard, k_0, k, is in general, accurate over a range of ground motion hazard curves and collapse fragility curves which are likely to occur in practise. In addition to the error associated with the power-model representation of ground motion hazard, the demand hazard closed-form solution also assumes a power-model for the median demand-intensity relationship and constant logarithmic standard deviation. These additional two assumptions will introduce further error in the demand hazard [12] in addition to the assumptions in the ground motion hazard.

4.9 Conclusions

This study has investigated the error associated with the assumptions necessary to obtain the closed form for the annual rate of structural collapse. The potential sources contributing to the error between the closed-form analytical solution and the exact solution for the annual rate of structural collapse were identified to be the curvature of the ground motion hazard and the dispersion in the collapse fragility curve, and the influence of these sources was been investigated via a parametric study. It was shown that the error in the closed-form solution is very sensitive to the log-log slope of the ground motion hazard curve, k, used, and while several fitting methods can be used to determine the value of k none are effective over the large range of likely values of parameters used. While the closed form analytical solution for the annual rate of structural collapse is without doubt insightful, considering that the numerical evaluation of the annual rate of collapse is straightforward, the authors recommend that future performance-based design methods should not consider use of the closed-form solution a necessity.

4.10 Acknowledgements

Financial support of the first author by the New Zealand Tertiary Education Commission is greatly appreciated.

4.11 References

[1] FEMA-273. NEHRP guidelines for the seismic rehabilitation of buildings. Federal Emergency Management Agency, Washington, DC, 1997.

[2] FEMA-302. NEHRP recommended provisions for seismic regulations for new buildings and other structures. Federal Emergency Management Agency, Washington, DC, 1997.

[3] FEMA-350. Recommended seismic design criteria for new steel momentframe buildings. SAC Joint Venture, Washington, DC, 2000.

[4] FEMA-356. Pre-standard and commentary for the seismic rehabilitation of building. Federal Emergency Management Agency, Washington, DC, 2000.

[5] SEAOC. Vision 2000: A Framework for Performance-based Design. Structural Engineers Association of California, Sacramento, 1995.

[6] Kennedy RP and Short SA. Basis for Seismic Provisions of DOE-STD-1020. Lawrence Livermore National Laboratory and Brookhaven National Laboratory, Washington, D.C., 1994.

[7] Cornell CA. Reliability-based earthquake-resistant design—the future, in 11th World Conference on Earthquake Engineering, Acapulco, Mexico, 1996.

[8] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[9] Cornell CA, Jalayer F, Hamburger RO, and Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. Journal of Structural Engineering 2002; 128(4): 526–533.

[10] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1901-1914.

[11] Mackie KR and Stojadinovic B. Performance-based seismic bridge design for damage and loss limit states. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1953-1971, DOI: 10.1002/eqe.699.

[12] Aslani H and Miranda E. Probability-based Seismic Response Analysis. Engineering Structures 2005; 27(8): 1151-1163.

[13] Bradley BA, Dhakal RP, Cubrinovski M, Mander JB, and MacRae GA. Improved seismic hazard model with application to probabilistic seismic demand analysis. Earthquake Engineering and Structural Dynamics 2007; 36(14): 2211-2225.

[14] Aslani H and Miranda E. Delivering improved information on seismic performance through loss deaggregation, in 8th National Conference on Earthquake Engineering, San Francisco, CA, 2006, paper no. 1126.

[15] Bazzurro P and Cornell CA. Disaggregation of seismic hazard. Bulletin of the Seismological Society of America 1999; 89(2): 501-520.

[16] Baker JW, Cornell CA, and Tothong P. Disaggregation of Seismic Drift Hazard, in 9th International Conference on Structural Safety and Reliability (ICOSSAR09), Rome, Italy, 2005.

[17] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[18] Jalayer F. Direct probabilistic seismic analysis: Implementing non-linear dynamic assessments. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2003, 173 pp.

[19] Vamvatsikos D and Cornell CA. Incremental dynamic analysis. Earthquake Engineering and Structural Dynamics 2002; 31(3): 491–514.

[20] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges. Engineering Structures 2007; 29(10): 2662-2672, DOI: 10.1016/j.engstrut.2006.12.015.

[21] Vamvatsikos D and Cornell CA. Applied incremental dynamic analysis. Earthquake Spectra 2004; 20(2): 523–553.

[22] Ibarra LF and Krawinkler H. Global collapse of frame structures under seismic excitations. University of California at Berkeley, Berkeley, CA, 2005.

[23] Stirling MW, McVerry GH, and Berryman KR. A new seismic hazard model for New Zealand. Bulletin of the Seismological Society of America 2002; 92(5): 1878–1903.

[24] Baker JW and Cornell CA. A vector-valued ground motion intensity measure consisting of spectral acceleration and eplison. Earthquake Engineering and Structural Dynamics 2005; 34(10): 1193-1217.

[25] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1973-1997, DOI: 10.1002/eqe.694.

[26] Luco N and Cornell CA. Seismic drift demands for two SMRF structures with brittle connections. Structural Engineering World Wide, 1998.

[27] Tothong P and Luco N. Probabilistic Seismic Demand Analysis Using Advanced Ground Motion Intensity Measures. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1837-1860.

[28] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthquake Spectra 2007; 23(2): 357-392, DOI: 10.1193/1.2723158.

[29] Kennedy RP, Cornell CA, Campbell RD, Kaplan S, and Perla HF. Probabilistic seismic safety study of an existing nuclear power plant. Nuclear Engineering and Design 1980; 59(2): 315–338.

[30] Haselton CB. Assessing Collapse Safety of Modern Reinforced Concrete Moment Frame Buidlings. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2007, 312.

5.Prediction of Spatially Distributed Seismic Demands in Specific Structures: Ground Motion and Structural Response

Bradley BA, Dhakal RP, MacRae GA, Cubrinovski M. Prediction of spatially distributed seismic demands in structures: ground motion and structural response. *Earthquake Engineering and Structural Dynamics* 2009. (submitted).

5.1 Abstract

The efficacy of various ground motion intensity measures (IM's) in the prediction of spatially distributed seismic demands (Engineering Demand Parameters, EDP's) within a structure is investigated. This has direct implications to building-specific seismic loss estimation, where the seismic demand on different components is dependent on the location of the component in the structure. Several common intensity measures are investigated in terms of their ability to predict the spatially distributed demands in a 10-storey office building, which is measured in terms of maximum interstorey drift ratios and maximum floor accelerations. It is found that the ability of an IM to efficiently predict a specific EDP depends on the similarity between the frequency range of the ground motion which controls the IM and that of the EDP. An IM's predictability has a direct effect on the median response demands for ground motions scaled to a specified probability of exceedance from a ground motion hazard curve. All of the IM's investigated were found to be insufficient with respect to at least one of magnitude, source-to-site distance, or epsilon when predicting all peak interstorey drifts and peak floor accelerations in a 10-storey RC frame structure. Careful ground motion selection and/or seismic demand modification is therefore required to predict such spatially distributed demands without significant bias.

5.2 Introduction

The seismic response of structural systems is complex with ground motion shaking causing a response that varies significantly in space and time. This complex response can result in significantly different acceleration and displacement demands at spatially different locations in a structure, which will depend on its dynamic characteristics, as well as the properties of the ground motion record exciting the structure.

In emerging performance-based frameworks such as the Pacific Earthquake Engineering Research (PEER) Centre performance-based earthquake engineering (PBEE) methodology [1] uncertainties in all aspects (from ground motion to loss estimation) of the seismic analysis of structures can be explicitly incorporated and propagated to obtain performance measures useful for decision making. In such a probabilistic framework, there are transparent advantages in being able to reduce uncertainties in each of the aforementioned aspects, since uncertainties inevitably result in an increase in the risk of structural failure and/or economic losses for infrequent hazards such as those posed by seismic-induced ground shaking.

An area of research in the past decade has been the investigation of ground motion intensity measures (IMs) which provide the link between the seismic hazard curve (which gives the probability/frequency of exceedance of a specific level of IM) and structural response (giving the distribution of the engineering demand parameter, EDP, for a given IM). An 'optimal' intensity measure must possess efficiency [2], sufficiency [3], predictability [4] and scaling robustness [5]. The aspects of efficiency and sufficiency have been studied in detail by Cornell and co-workers (e.g. [5-9]) where the seismic response of structures was measured simply via the maximum interstorey drift over all floors (which relates well to joint rotations in structural elements and therefore the potential for structural collapse). Predictability relates to the accuracy in predicting an IM from ground motion prediction equations. With the increased interest in ground motion selection methods (e.g. [10]), scaling robustness seeks to determine if the distribution of EDP using scaled ground motions is biased compared with that obtained using un-scaled ground motions [11]. Optimal intensity measures for total floor accelerations have received less attention than that of peak interstorey drifts, with the exception of Taghavi and Miranda [12] who examined the efficiency of four different IMs at predicting peak floor accelerations using simple elastic structural models.

The significant spatial variation in the response of structural systems with several or more storeys means that separate consideration must be given to each of these demands when

rigorously considering the seismic performance of such systems within the PEER framework. As such, loss estimation methods used within the PEER PBEE framework typically employ a vector of EDP's which account for these spatially varying demands (typically maximum interstorey drift ratios and maximum floor accelerations). Furthermore, ground motion IMs which are efficient, sufficient, predictable and robust to scaling for this vector of EDP values are required.

Aslani [13] considered the efficiency and sufficiency of four different IM's for use in predicting spatially distributed demands in structures, and this research is intended to extend the work of Aslani [13] in the following ways: (1) seismic hazard curves for each of the ground motion IM's are developed independently allowing explicit consideration of the predictability of the different IM's; (2) consideration is given to efficiency, sufficiency, predictability, and scaling robustness of the IMs; (3) ground motion selection based on hazard deaggregation is employed; and (4) 50^{th} percentile rotation independent geometric mean (GMRotI50) intensity measures are used in both hazard computations and seismic response analysis.

The purpose of this chapter is to use common IM's presented in literature, and for a specific structure investigate their predictability, efficiency, sufficiency, and scaling robustness in predicting peak interstorey drifts and total floor acceleration demands throughout the structure. In a companion chapter [14] correction for IM insufficiency, as well as the resulting demand hazard, collapse hazard, and loss estimation results based on the different IM's is considered.

5.3 Structure considered

The case study structure used herein is based on the geometry of the 'Red Book building' [15], a ten storey reinforced concrete (RC) structure, which acts as a design example of the New Zealand Concrete Code [16]. The primary lateral load carrying system consists of four one-way perimeter moment resisting frames which are 3 bays long. Vertical loads are transferred primarily through interior columns with gravity beams supporting one-way floor units.

A perimeter frame 2D model was developed using the finite element analysis program OpenSees [17]. Due to the symmetry of the structure, it was assumed that the 3D response could be reasonably approximated by separate 2D analyses in each of the two primary frame directions. The effects of foundation flexibility due to soil-foundation interaction were

considered simply by using elastic rotational springs at the base of the columns [18]. The structure was modelled using a lumped mass model and non-linear (beam) elements with the appropriate backbone properties determined using fibre-based section modelling, and stiffness and strength degradation based on calibration with experimental tests [19]. The structural model had a fundamental period of T_1 = 1.74 seconds. Based on a pushover analysis it was determined that the 'yield' displacement, d_y, of the structure was 10 cm (this is used for the inelastic spectral displacement IM, S_{di}).

The seismic demand due to ground motion excitation was monitored via peak interstorey drift ratios and peak floor accelerations at each floor in the structure (i.e. a total of ten drifts and eleven accelerations).

5.4 Ground motion intensity measures and seismic hazard

In order to investigate the prediction of spatially distributed demands in structures a variety of ground motion IMs are selected. As there have been numerous ground motion IMs presented in the literature relating to various different aspects of structural behaviour it is necessary to apply some criteria to determine which IMs to investigate in this research. Firstly, it was desired to consider several IMs which have been used by other researchers when examining structural response from a probabilistic viewpoint. Secondly, and more importantly, all IMs used had to have a 'robust' ground motion prediction equation which can be used to develop seismic hazard curves using this IM at a variety of sites. This second point is particularly important as many studies have focused on the consideration of somewhat complex IMs which may be a combination of several 'standard' IMs in an effort to achieve better response prediction (i.e. efficiency). However, without a ground motion prediction equation for such an IM, no ground motion hazard curves can be developed, and hence no PBEE assessment can be performed using this IM. The term 'robust' has been used in order to exclude IMs which have simple ground motion prediction equations based on limited data and applicability to various sites, from comprehensive ground motion prediction equations based on large ground motion databases and considering many features which affect ground motion prediction (e.g. faulting types, hanging wall effects, local soil effects) such as those of the next generation attenuation (NGA) project [20].

Based on the above criteria a total of five different ground motion IMs were selected, namely: peak ground acceleration (*PGA*); peak ground velocity (*PGV*), elastic spectral displacement (S_{de}); inelastic spectral displacement (S_{di}); and spectrum intensity (*SI*). *PGA* and

S_{de} can be predicted from (elastic) spectral acceleration prediction equations, some of which now also include coefficients for computing *PGV* [20]. Prediction of S_{di}, defined as the peak displacement of a bilinear single degree-of-freedom oscillator [21], is obtained by combining a ground motion prediction equation for S_{de} with a ground motion prediction for the ratio S_{di}/S_{de} [21] (while the empirical equation for S_{di}/S_{de} was determined using one specific ground motion prediction equation Tothong and Cornell [21] argue that it can be used with any prediction equation for S_{de}). Finally, a ground motion prediction equation for *SI*, defined as the integral of the pseudo-spectral velocity from 0.1-2.5 seconds [22], can be computed directly from ground motion prediction equations for spectral acceleration [23]. The effect of epistemic uncertainties is beyond the scope of this study. However, it should be noted that as all ground motion IM's have 'robust' prediction equations, then the effect of epistemic uncertainty is expected to be of a similar magnitude for all the different IMs.

A simple hypothetical site with a 30-m averaged shear wave velocity of 600 m/s was considered as illustrated in Figure 5-1a, which is a closest distance of 15 km from a 40 km strike-slip fault. The fault has a Gutenberg-Richter magnitude distribution with $\alpha = 3.0$ and $\beta = 0.8$; minimum and maximum magnitudes of 5.0 and 7.5, respectively; and events assumed to be Poissonian in time. Based on this hypothetical scenario and using the Boore and Atkinson [24] ground motion prediction equation for *PGA*, *PGV* and S_{de}; the prediction equation of Tothong and Cornell [21] for S_{di}; and the prediction equation of Bradley *et al.* [23] for *SI* (with Boore and Atkinson [24] used as the 'base' prediction equation for both S_{di} and *SI*), the ground motion hazard curves shown in Figures 5-1b-5-1d were determined using the probability-based formulation for the ground motion hazard [25]. Because of the initial period of the structure it is observed that the elastic and in-elastic spectral displacement hazards are similar in Figure 5-1c (i.e. the equal displacement rule). Note that the 50th percentile orientation-independent geometric mean of the two horizontal ground motion components, GMRotI50 [26], definition for the ground motion intensity measures has been used in computing the hazard curves for all five IMs.

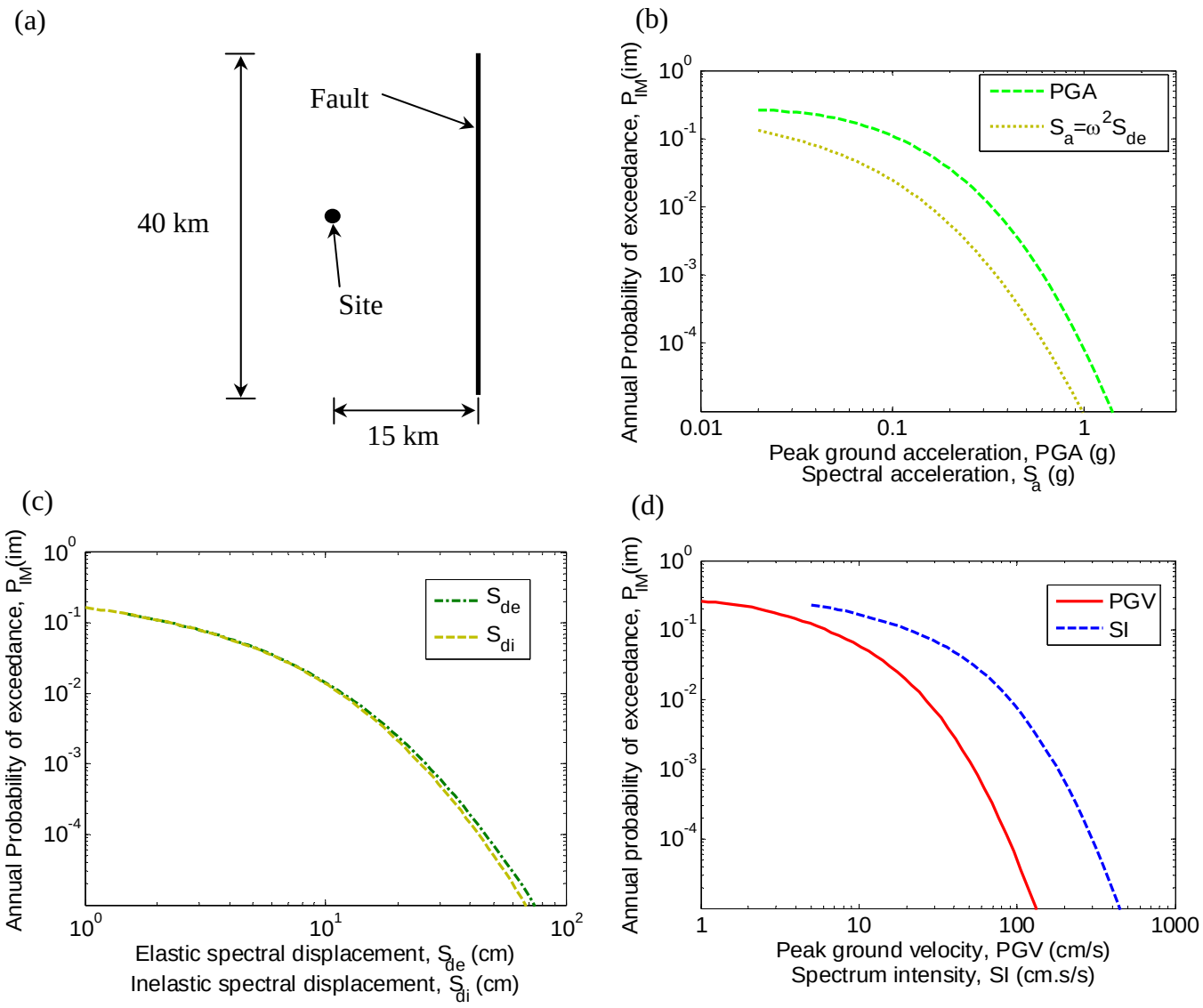


Figure 5-1: Hypothetical site considered and ground motion hazard curves for the five different ground motion intensity measures (IMs) considered.

5.5 Ground motion selection

It is important to carefully select ground motion records for use in dynamic analysis to avoid bias in structural response [27]. Ground motion records should be selected which are representative of those which are most likely to occur at the site of interest in the future. Hence, the target is to select a suite of ground motions which has the same distribution of ground motion properties as the deaggregation [28, 29] of the seismic hazard at the site. In the remainder of the chapter particular attention will be given to the spatial distribution of seismic demands for ground motions which have exceedance probabilities of 1/475 and 1/2475 (corresponding to 10% and 2% probabilities of exceedance in 50 years, respectively). Table 5-1 gives the mean moment magnitude (M_w), (Boore-Joyner) source-to-site distance

(R), and epsilon (ε, defined as the number of standard deviations a specific ground motion parameter is above the predicted mean) for the five different IMs obtained by deaggregation of the seismic hazard at these exceedance probabilities. It can be seen that for a given IM, the mean M_w, and ε are larger and mean R smaller for the 1/2745 exceedance probability compared to the 1/475 exceedance probability. In addition, for a given exceedance probability it is observed that there are quite significant differences between the mean M_w, R and ε values for the different IMs, since they are affected by different properties of the ground motion. M_w and R obviously affect the intensity, frequency content and duration of ground motion records, and it has also been shown that when spectral ordinates are used as a ground motion IM (i.e. PGA and S_{de} used in this study) that ε has an effect on structural response, as it relates to spectral shape [6]. S_{di} has been shown to be (relatively) insensitive to ε (when predicting the peak interstorey drift over all floors), since S_{di} directly accounts for spectral shape in the case of period elongation [5]. The effect of ε on PGV and SI has not been researched in detail.

Table 5-1: Statistics of the ground motion hazard deaggregation and suite of ground motion records used.

	$P_{IM}(im) = 1/475$				$P_{IM}(im) = 1/2475$				Ground motion suite		
	IM	μ_{M_w}	μ_R	μ_ε	IM	μ_{M_w}	μ_R	μ_ε	μ_{M_w}	μ_R	μ_ε
PGA (g)	0.515	6.93	18.8	1.58	0.744	7.01	18.5	2.14	6.49	20.8	1.17
PGV (cm/s)	44.1	7.16	18.8	1.26	66.2	7.24	18.5	1.88	6.49	20.8	1.58
S_{de} (cm)	20.9	7.04	18.9	1.46	33.4	7.10	18.7	2.06	6.49	20.8	1.57
S_{di} (cm)	20.1	7.03	19.0	1.44	31.6	7.10	18.8	2.08	6.49	20.8	1.57
SI (cm.s/s)	148.3	7.11	18.7	1.34	223.3	7.17	18.3	1.89	6.49	20.8	1.76

As the value of ε depends on the IM and GMPE used and cannot currently be determined *a priori* (using, for example, the PEER NGA ground motion database [30]) then using the deaggregation results in Table 5-1, ground motions were initially selected based on wide range of M_w, R, and site 30-m averaged shear-wave velocity (Vs_{30}) (specifically $6.0 < M_w < 8.0$; $0 < R < 30$ km; $300 < Vs_{30} < 800$ m/s) giving a total of 155 ground motions (each with two orthogonal horizontal components). The ε values for the five different ground motion parameters were determined and then the allowable ranges of M_w, R, Vs_{30}, and ε were further constrained to obtain a set of 25 ground motions (i.e. 50 different horizontal ground motion records for use in dynamic analysis) which were based on $6.2 < M_w < 7.7$; $10 < R < 28$ km; $300 < Vs_{30} < 800$ m/s; ε (of any IM) > 0.4. No constraints were placed on the number of recordings from a single event. The adopted ground motion records are presented in Table

5-2 and their statistics for the different IMs are summarised in Table 5-1. Note that the process for selecting ground motion records was conducted in an iterative fashion as it was not possible to find a large enough suite of records which match the statistics (mean and standard deviation) of the seismic hazard deaggregation exactly. Therefore a trade-off was required to try and match all M_w, R, and ε values relatively well.

5.6 Prediction of structural response: Deterministic hazard scenario

Before investigating the efficacy of the five different ground motion IMs in predicting the spatially varying structural response for the 1/475 and 1/2475 probability of exceedance hazard levels, it is first necessary to test for any bias when comparing the different ground motion IMs. Potential sources of bias could be whether the selected ground motion record suite is equally representative for all five different ground motion IMs or whether the different ground motion prediction equations used to determine the scale factors (used to scale the amplitude of motion only) for each ground motion are consistent. To investigate the above points a deterministic earthquake scenario is considered with a moment magnitude of $M_w = 7.0$ and a source-to-site distance of $R = 18$ km. Note that this deterministic scenario is intentionally similar to the mean magnitude and distance obtained from deaggregation of the seismic hazard in Table 5-1. Table 5-3 gives the median and dispersion (lognormal standard deviation) for the ground motion IMs obtained from the ground motion prediction equations for the deterministic scenario. Figures 5-2a and 5-2b illustrate the median (specifically, the mean of the logarithms which is the median assuming a lognormal distribution) and dispersion for the GMRotI50 acceleration response spectra obtained by scaling the ground motion records to the median IM for the deterministic scenario. Since the median response spectra of the 25 ground motion records scaled based on the five different IMs are very similar, it indicates that the selected ground motion suite and the different ground motion prediction equations do not introduce any significant bias when comparing the results of the structural analyses to follow. Note that this result was to be expected since all of the IMs use (or are derived from) the Boore and Atkinson [24] ground motion prediction equation and ground motions were selected to match hazard deaggregation. Figure 5-2b provides insight into the effect of IM scaling on the dispersion in response spectra amplitudes as a function of vibration period. Obviously, *PGA* scales all the ground motions to have the same spectral acceleration at $T = 0$ so the dispersion is zero at $T = 0$, and similarly for S_{de} at $T = T_1$. Also since the

inelastic spectral displacement for this scenario, S_{di} = 6.09 cm, is less than the yield displacement of the inelastic single-degree-of-freedom (SDOF), d_y = 10 cm, then the dispersion is also zero at $T = T_1$ for ground motions scaled to S_{di} (i.e. the S_{de} and S_{di} lines in Figure 5-2b are coincident). An increase in response spectra dispersion with period is observed for ground motions scaled based on *PGA*, while in a similar fashion the dispersion increases as T moves away from T_1 for S_{de} - and S_{di} –based scaling. While scaling based on *PGV* or *SI* does not 'fix' the dispersion to zero at any point of the response spectra it is observed that apart from the small period window around $T = T_1$ (approximately $1.0 < T < 2.5$) the dispersion for both *PGV* and *SI* is lower than when ground motions are scaled to S_{de}.

Table 5-2: Properties of the ground motions adopted.

ID*	Earthquake	Year	Recording station	M_W	R (km)	Vs_{30} (m/s)
125	Friuli, Italy	1976	Tolmezzo	6.50	14.97	425
265	Victoria, Mexico	1980	Cerro Prieto	6.33	13.80	660
339	Coalinga	1983	Parkfield - Fault Zone 15	6.36	28.00	376
359	Coalinga	1983	Parkfield - Vineyard Cany 1E	6.36	24.83	339
369	Coalinga	1983	Slack Canyon	6.36	25.98	685
587	New Zealand	1987	Matahina Dam	6.60	16.09	425
755	Loma Prieta	1989	Coyote Lake Dam (SW Abut)	6.93	19.97	598
776	Loma Prieta	1989	Hollister - South & Pine	6.93	27.67	371
952	Northridge	1994	Beverly Hills - 12520 Mulhol	6.69	12.39	546
963	Northridge	1994	Castaic - Old Ridge Route	6.69	20.10	450
995	Northridge	1994	LA - Hollywood Stor FF	6.69	19.73	316
1003	Northridge	1994	LA - Saturn St	6.69	21.17	309
1010	Northridge	1994	LA - Wadsworth VA Hospital South	6.69	14.55	414
1077	Northridge	1994	Santa Monica City Hall	6.69	17.28	336
1485	Chi-Chi, Taiwan	1999	TCU045	7.62	26.00	705
2461	Chi-Chi, Taiwan	1999	CHY028	6.20	23.44	543
2495	Chi-Chi, Taiwan	1999	CHY080	6.20	21.34	553
2618	Chi-Chi, Taiwan	1999	TCU065	6.20	25.17	306
2619	Chi-Chi, Taiwan	1999	TCU067	6.20	27.66	434
2626	Chi-Chi, Taiwan	1999	TCU075	6.20	18.47	573
2627	Chi-Chi, Taiwan	1999	TCU076	6.20	13.04	615
2655	Chi-Chi, Taiwan	1999	TCU122	6.20	18.10	475
2661	Chi-Chi, Taiwan	1999	TCU138	6.20	21.11	653
3300	Chi-Chi, Taiwan	1999	CHY074	6.30	27.57	553
3507	Chi-Chi, Taiwan	1999	TCU129	6.30	22.69	664

*ID as given on the NGA database. http://peer.berkeley.edu/nga/earthquakes.html

Table 5-3: Median and dispersion in the ground motion IMs for the deterministic M_w = 7.0, R = 18 km scenario.

IM	Median, $exp(\mu_{lnIM})$	Dispersion, σ_{lnIM}
PGA (g)	0.187	0.564
PGV (cm/s)	15.78	0.560
S_{de} (cm)	6.19	0.690
S_{di} (cm)	6.09	0.672
SI (cm.s/s)	52.4	0.592

Figures 5-3a and 5-3b illustrate the median response (in terms of maximum interstorey drift ratios and maximum floor accelerations) of the case study structure based on ground motion scaling using the five different IMs. As expected, the median response, both for interstorey drifts and floor accelerations are approximately the same for all five IMs. The minor exception being that ground motions scaled using *PGA* give slightly larger interstorey drifts and floor accelerations over the lower portion of the structure. This is consistent with the slightly larger median response spectra over $1.0 < T < 3.0$s using *PGA*-scaling in Figure 5-2a. Figure 5-3c illustrates the dispersion in the maximum interstorey drifts for the deterministic scenario. It can be seen that over the lower half of the structure where the peak responses are primarily due to the first mode of vibration, that S_{de} (and S_{di}) are the most efficient in predicting the interstorey drifts, while *PGA* is the worst, and *SI* is marginally better than *PGV* as it contains spectral velocity information at periods around that of the first mode. In the upper-half of the structure, where the effects of higher vibration modes are more significant it is clearly seen that the efficiency of the spectral displacement IMs (S_{de} and S_{di}) reduces and the efficiency of *PGA* increases. Figure 5-3d illustrates the dispersion in the peak floor accelerations for the deterministic scenario. It is apparent that *PGA* has the best efficiency for all floors, although the difference is less pronounced in the upper floors (where the structures dynamic characteristics have significantly modified the ground motion input at the base). On the other hand, spectral displacement IMs (S_{de} and S_{di}) are the worst at predicting the maximum floor accelerations over all floors. Note that the dispersion in the ground floor peak acceleration when using *PGA* is not zero as the ground motion IMs are based on GMRotI50 [26] (i.e. the geometric mean of the two components), while only a single component is applied in each of the structural analyses. It is necessary to use the GMRotI50 definition for the IMs to be consistent with the ground motion hazard curves [31].

The results of Figures 5-3c and 5-3d are consistent with the results obtained by Aslani [13] and Taghavi and Miranda [12], and clearly indicate that the ability of various ground motion IMs to predict structural response EDPs depends on the similarity of the frequency

range of the motion which dominates the EDP and that of the IM. For example, peak accelerations are dominated by high frequency content so *PGA* is the most efficient IM, while displacements are dominated by lower frequency content so IMs in the lower frequency region (i.e. S_{de} and S_{di} in this case) are more efficient in displacement prediction. The velocity IMs (*PGV* and *SI*) being in the 'medium' frequency range (i.e. between accelerations and displacements) provide 'moderate' efficiency in predicting both maximum interstorey drifts and maximum floor accelerations.

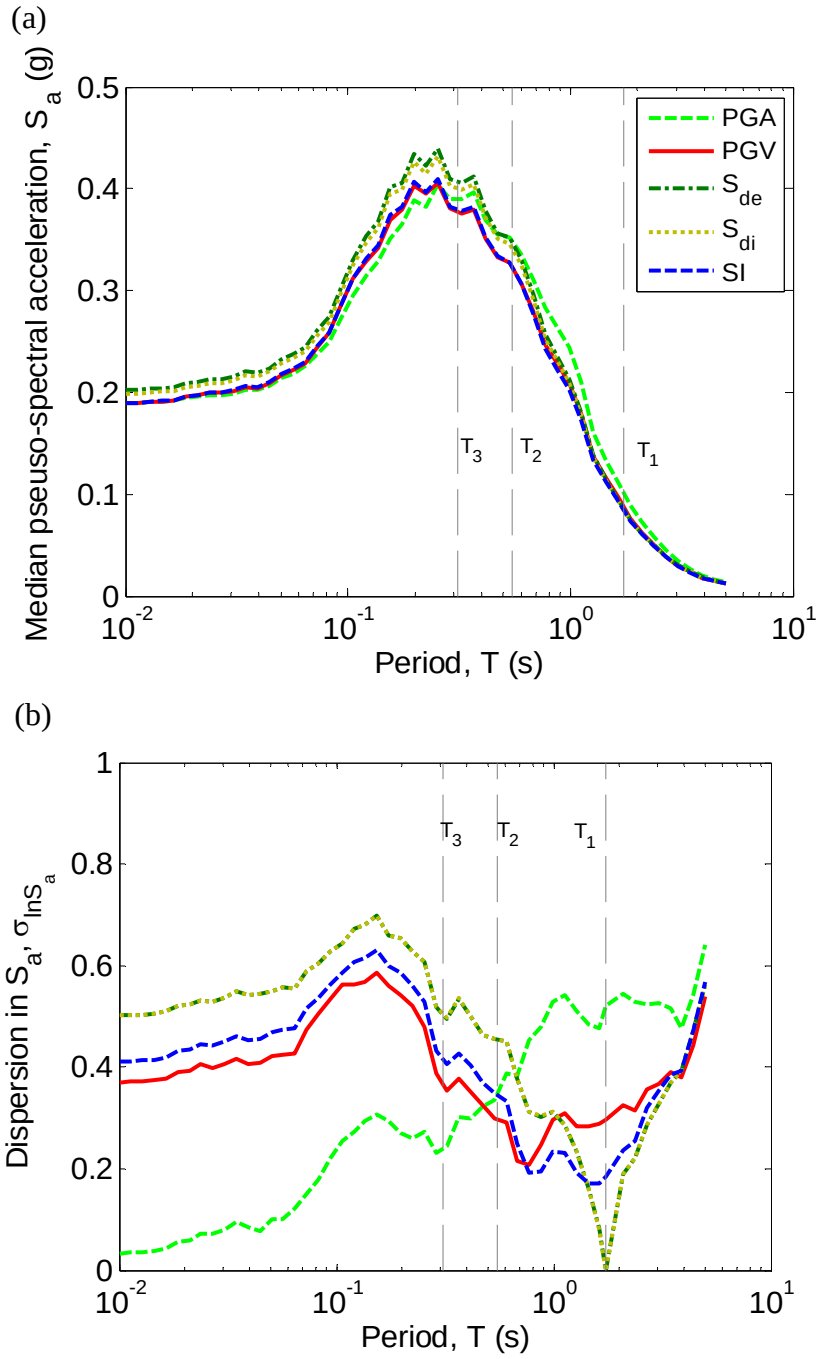


Figure 5-2: Response spectra of the ground motion suites scaled based on the different IMs for a M_w = 7, R = 18 km scenario

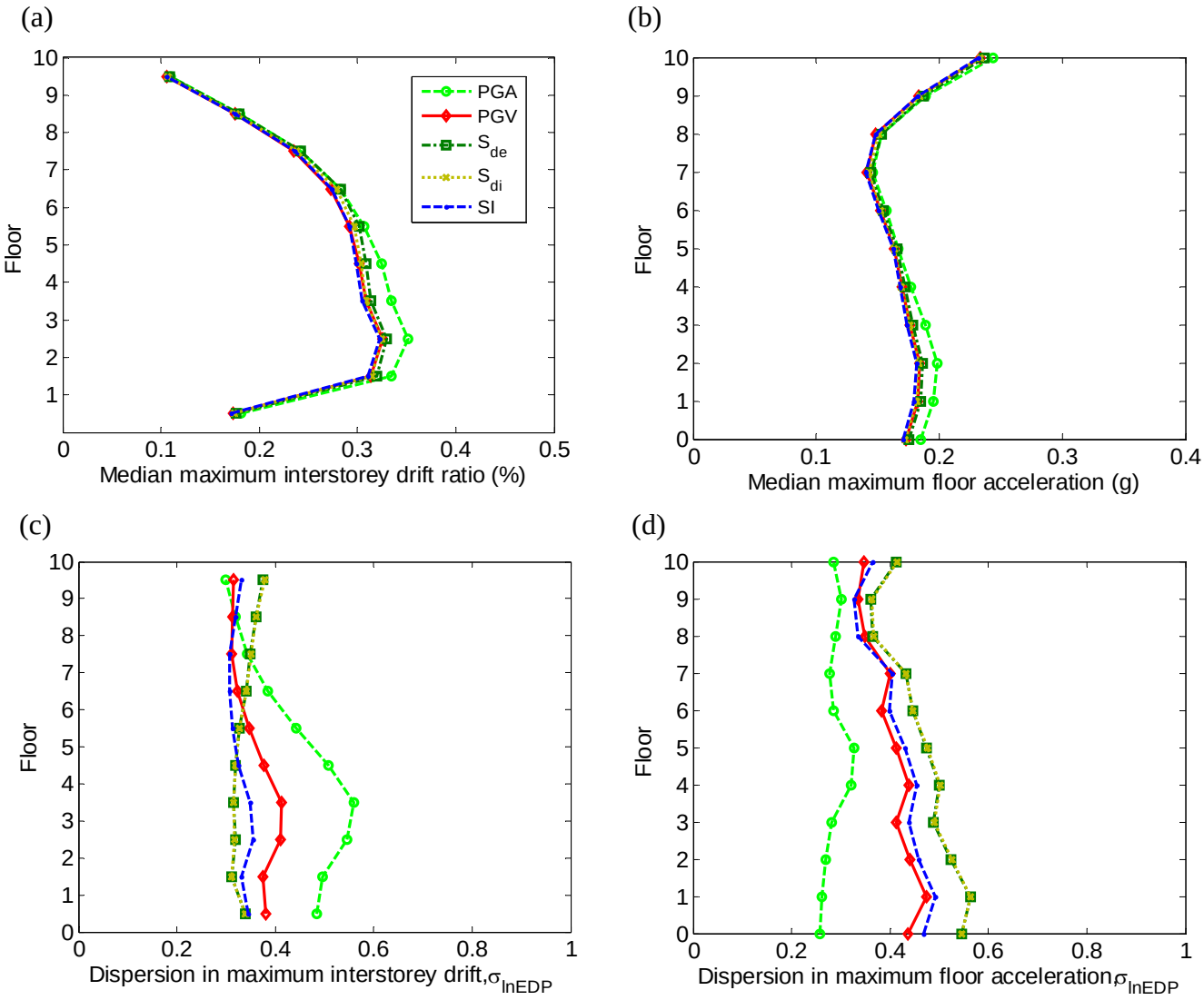


Figure 5-3: Median and dispersion in maximum interstorey drift and maximum floor acceleration demands for the various IMs for the deterministic scenario.

5.7 Prediction of structural response: Probabilistic hazard

The nature of the deterministic scenario presented in the previous section allowed Figures 5-3c and 5-3d to be viewed solely to investigate the efficiency of the five different IMs (in this case for a relatively small level of ground motion shaking). As mentioned previously however, the determination of an optimal IM contains several other criteria, one of which is the predictability of the IM. Predictability relates to the magnitude of the aleatory uncertainty in the ground motion prediction equation used to compute the ground motion hazard for a specific site. Predictability is an important property of a ground motion IM since it will affect the probability of a specific level of ground motion occurring. Bommer and Abrahamson [32, Figure 3] illustrate that the effect of a large uncertainty in a ground motion

prediction equation (i.e. poor predictability) is to increase the likelihood of a specific level of ground motion intensity occurring, with the increase in likelihood becoming more significant at long return periods. As a direct indication of the predictability of the five different IMs considered here, Table 5-3 indicates the dispersion in the ground motion prediction equations for the deterministic scenario considered in the previous section. As is typical, the standard deviation of a response spectral ground motion prediction equation increases with response spectral period, as long period motion is more deterministically related to the earthquake source [33, p101] than short period motion (i.e. the predictability of S_{de} and S_{di} is worse than *PGA*), while the predictability of *PGA*, *PGV* and SI are similar.

5.7.1 Interstorey drift response

To illustrate the effects of predictability on the results of structural response analyses the ground motion records given in Table 5-2 were scaled to ground motion intensities which had 1/475 and 1/2475 exceedance probabilities (using the hazard curves in Figure 5-1), the values for which are given in Table 5-1. Figure 5-4 illustrates the median and dispersion in the maximum interstorey drifts predicted using the various IMs for the two different exceedance probabilities. Unlike the deterministic scenario where all of the five IMs produced similar median demands, it is clear from Figures 5-4a and 5-4b that there is a significant difference when the ground motions are scaled to the same exceedance probabilities. For example, scaling ground motions to *SI* gives median values for the maximum interstorey drift between the 2nd and 3rd floors of 0.9% and 1.25% at the two different exceedance probabilities compared to 1.1% and 1.6% using ground motions scaled to S_{de} (i.e. 22% and 28% differences, respectively). It should be clear that the relative magnitude of the median values of the maximum interstorey drifts between the different IMs is closely related to the predictability of the different IMs. Figures 5-4c and 5-4d illustrate the dispersion in the maximum interstorey drifts obtained using the five different IMs for the 1/475 and 1/2475 exceedance probabilities. The trends regarding efficiency in predicting maximum interstorey drifts are similar to those for the deterministic scenario with the key difference being that as the intensity of the ground motion increases inelastic response causes changes in the vibration characteristics of the structure. These changing vibration characteristics subsequently affect the efficiency in predicting the EDPs at different locations in the structure. For example, comparing the dispersion in the interstorey drifts at the 1/475 and 1/2475 exceedance probabilities illustrates that increasing inelastic behaviour reduces the dispersion in the

prediction of the interstorey drifts in the lower half of the structure and increases the dispersion in the upper floors.

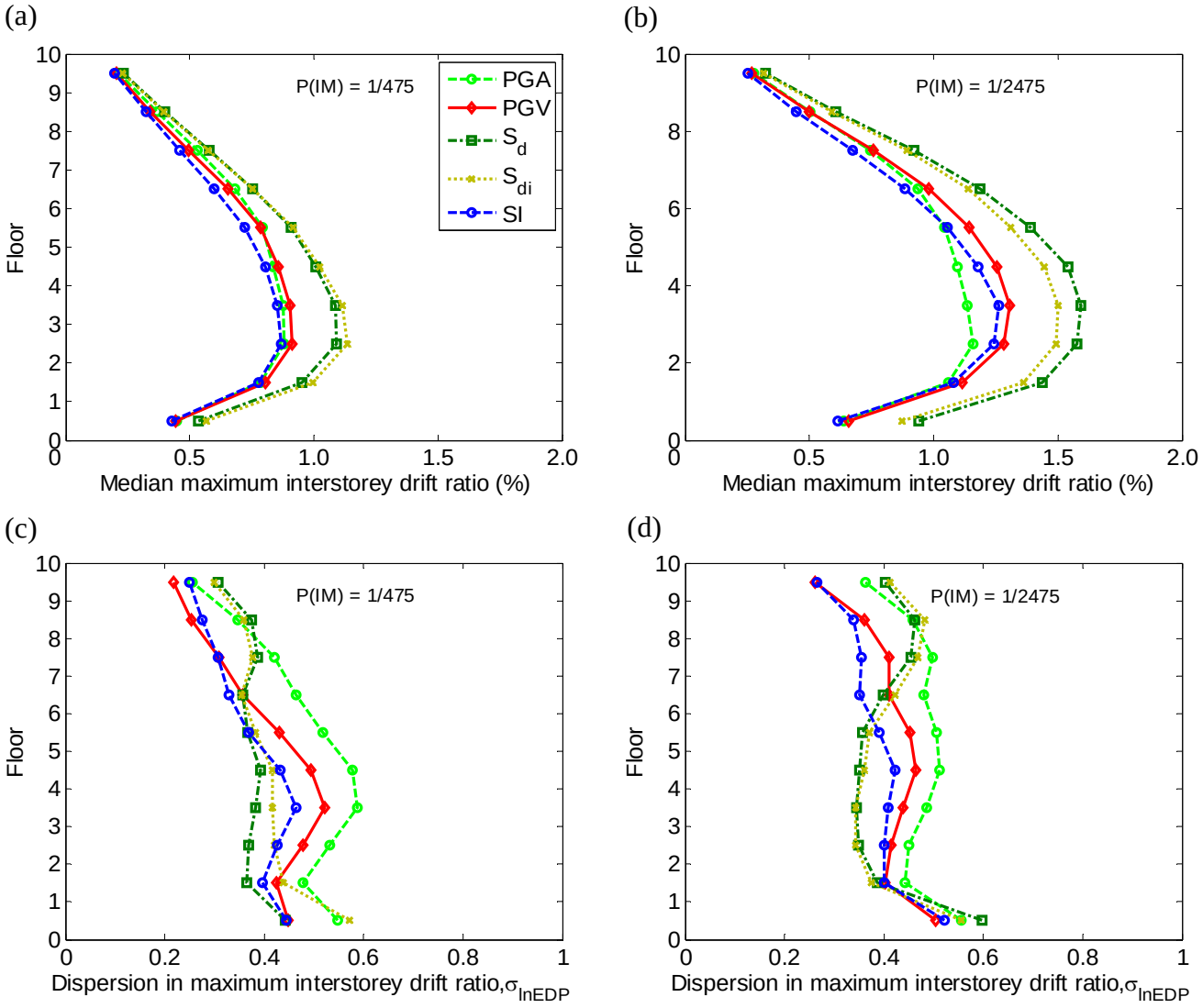


Figure 5-4: Median and dispersion is the maximum interstorey drifts for ground motions scaled to the 1/475 and 1/2475 exceedance probabilities using the various IMs.

5.7.2 Floor acceleration response

Figures 5-5a and 5-5b illustrate the median values of the maximum floor accelerations for the various IMs. Similar to the median values of the interstorey drifts, it is observed that the relative magnitude of the maximum floor accelerations is directly related to the predictability of the IMs. A comparison of Figures 5-3b, 5-5a and 5-5b illustrates the change in the spatial distribution of the peak floor accelerations as the ground motion intensity increases. For the aforementioned deterministic scenario the largest floor accelerations occur at the top of the structure due to the presence of significant higher mode effects. However as

the ground motion intensity increases, inelastic behaviour in the lower floors of the structure (which causes an elongation in the effective period of the structure) effectively acts as a filter on the high frequency components of the ground motion. The same logic also explains why the maximum interstorey drift demands in the upper floors of the structure reduce (relative to the maximum interstorey drifts in the lower floors) as the ground motion intensity increases. Figures 5-5c and 5-5d illustrate the dispersion in the prediction of the maximum floor accelerations at the 1/475 and 1/2475 exceedance probabilities. Similar trends are observed compared to the deterministic scenario with *PGA* being the most efficient IM and spectral displacements the worst. However, due to significant inelastic behaviour at the 1/2475 exceedance probability, it is seen that the efficiency of *PGA* in predicting peak accelerations on some floors is reduced.

It is interesting to observe in Figures 5-4 and 5-5 that there is little difference between the predictive capacity of S_{di} and S_{de} for the structure considered. Comparison with the results of Tothong and Luco [5, Table 2] however illustrates that S_{di} provides little improvement over using S_{de} for structures with fundamental period above 1.5 seconds. Also as careful ground motion selection has been used with respect to ε, then the results presented here for S_{de} are equivalent to somewhere between S_{de} and S_{de}&ε presented in Tothong and Luco [5, Table 2]. In such cases an intensity measure, $IM_{1I\&2E}$, which combines information of the inelastic first mode response and elastic higher modes can give better approximations of drift demands [5]. However, as $IM_{1I\&2E}$ combines different mode contributions in terms of spectral displacements, then those spectral displacements due to higher modes are relatively small and it is not likely that $IM_{1I\&2E}$ will be efficient in predicting floor acceleration demands. $IM_{1I\&2E}$ was not considered here as an IM since current ground motion prediction equations for $IM_{1I\&2E}$ require the use of a first-order Taylor-series expansion (which is known to be inaccurate for the large uncertainty in ground motion prediction equations), and the 'equal displacement' assumption [34].

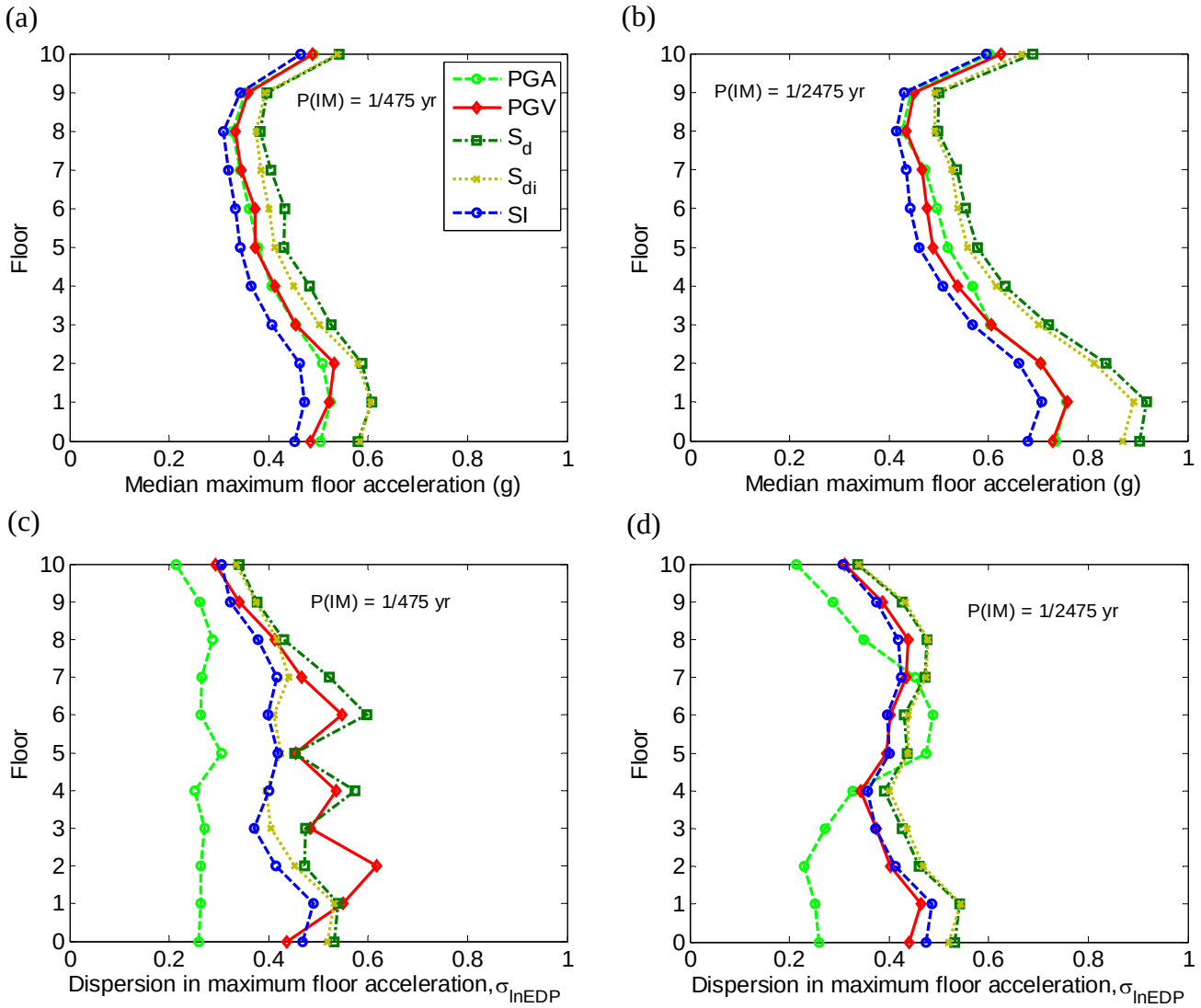


Figure 5-5: Median and dispersion in the maximum floor accelerations for ground motions scaled to the 1/475 and 1/2475 exceedance probabilities using the various IMs.

Figures 5-4a, 5-4b, 5-5a and 5-5b have illustrated that for a given exceedance probability, a significant reduction in the median response (specifically for this structure), both interstorey drifts and floor accelerations, can be obtained by using a predictable intensity measure such as *PGA*, *PGV*, and *SI*. This has a direct implication for current code-based applications for time-history analysis which require that the average response be used for design if seven or more ground motion records are used [35]. Note that design codes state that the 'average' of the structural responses from the different ground motions should be used, where it is assumed that 'average' refers to the arithmetic mean of the responses (and not the mean of the logarithms of the responses used here to get the 'median'). Since for a lognormal distribution the ratio of the mean to the median is $\exp(\sigma^2_{\ln X}/2)$, where $\sigma_{\ln X}$ is the dispersion, then it was found that *PGA*, *PGV* and *SI* give lower 'mean' responses than spectral displacements, S_{de} and S_{di}.

5.7.3 Conditional distribution of seismic demand given intensity measure

In order to proceed from seismic response analysis results and compute demand hazard and loss estimation it is necessary to know the conditional distribution of the demand (*EDP*) given ground motion intensity (IM), i.e. $f(EDP|IM)$. Due to the numerous permutations of the 21 different EDPs, five IMs and two different hazard levels considered, no attempt is made here to rigorously illustrate the observed distribution although it is pertinent to discuss such a topic. Reference is given to the numerous studies which have found that interstorey drifts [e.g., 2, 36] and floor accelerations [e.g., 36] are lognormally distributed. Note that Taghavi and Miranda [12] argue that peak floor accelerations are normally distributed, however their comparisons between empirical and analytic distributions were for ground motion IMs with dispersions which were less than 0.3 (in which case the normal and lognormal distributions are similar). Although not explicitly illustrated here, it was found that the conditional demand distribution, $f(EDP|IM)$ was satisfactorily estimated using the lognormal distribution based on the Kolmogorov-Smirnov goodness-of-fit test [37] at a 5% significance level (i.e. α=0.05) for all of the *EDP*|*IM* permutations (which as shown in Figures 5-3-5-5 included dispersions from 0.25-0.75).

5.8 Sufficiency of intensity measures

The previous two sections have focused on the efficiency and predictability of an IM for estimating the spatially distributed demands in structures. The remaining property of an

optimal IM is sufficiency; scaling robustness [5] can be thought of as sufficiency with respect to scale factor. Sufficiency (in this context) relates to the conditional independence of the distribution of *EDP* given *IM* on other parameters such as earthquake magnitude (M_w), source-to-site distance (R), and epsilon (ε). IM sufficiency is important since the distribution, $f(EDP|IM)$ is obtained from the results of a finite number of seismic response analyses. Thus if the distribution $f(EDP|IM)$ is dependent on the M_w, R and ε values of the ground motions used, then the distribution will be biased if the distribution of M_w, R and ε of the ground motion records used in the seismic response analysis is not the same as that of the ground motions which will occur at the site in the future. Thus, mathematically speaking, sufficiency requires that $f(EDP|IM) \cong f(EDP|IM, M_w, R, \varepsilon)$, where the 'approximately equals' sign is intended to mean 'practically equal to'. In order to test for sufficiency linear regression is typically performed between some property of the ground motion records (i.e. M_w, R, ε) and the observed EDPs from the seismic response analysis [3]. Thus the coefficient b from the linear regression line $E[\ln EDP] = a + bx$ (where x is one of M_w, R, ε) indicates the dependence of the observed EDP values on the parameter x. Since the linear regression is based on a finite number of observations it is necessary to use statistical tests to determine the significance of the coefficient b. As an underlying assumption of linear regression is that the observations, lnEDP, are normally distributed, and since there are a finite number of observations, it follows that the coefficient, b has a student-t distribution and the F-test can be used to determine the statistical significance of b [37]. The F-test gives a 'p-value' which is the probability of the coefficient b having a value at least as large as that observed, given that its underlying true value is zero. Typically p-values less than 0.05 are used to indicate a statistically significant value. For example, Figure 5-6a illustrates the sufficiency of *PGV* with respect to source-to-site distance, R, for the peak interstorey drift between the 2nd and 3rd floors, while Figure 5-6b illustrates the sufficiency of S_{de} with respect to epsilon for the peak 2nd floor acceleration (both for the 1/475 exceedance probability). Figure 5-6a indicates that based on the finite observations the positive correlation between ln*EDP* and R is statistically insignificant (p-value = 0.12 > 0.05), while the negative correlation between ln*EDP* and ε in Figure 5-6b is statistically significant (p-value = 2.4×10^{-6} << 0.05). Because of the numerous permutations for the five different IMs over the 21 different EDPs monitored in the case study structure, discussion in the remainder of the section is given with respect to p-values directly.

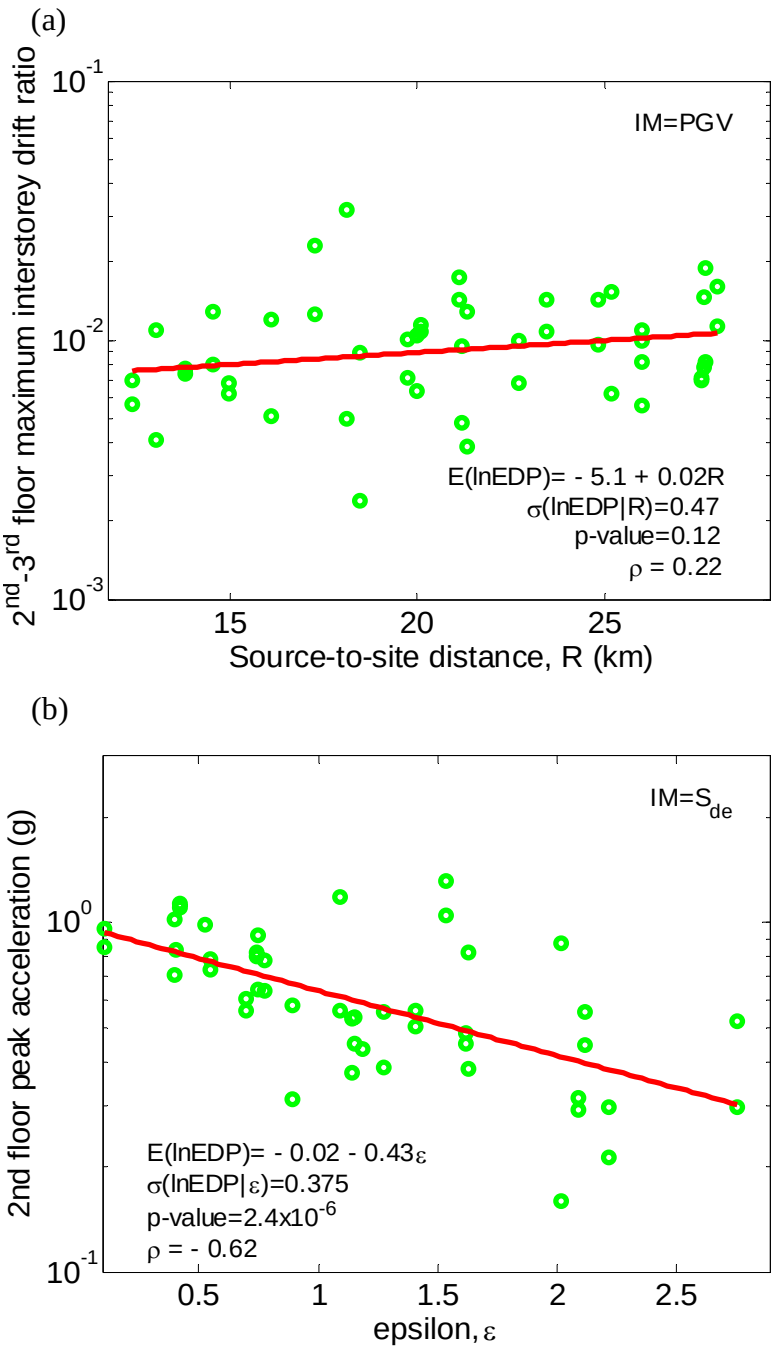


Figure 5-6: Sufficiency of: (a) peak ground velocity (*PGV*) with respect to source-to-site distance in predicting the peak interstorey drift ratio between the 2nd and 3rd floors; and (b) spectral displacement (S_{de}) with respect epsilon in predicting the 2nd floor peak acceleration.

5.8.1 Sufficiency with respect to magnitude and source-to-site distance

Several of the ground motion IMs used here have been investigated (regarding M_w and R sufficiency) previously for the peak interstorey drift and floor accelerations over all floors and are briefly reviewed here. Aslani [13] showed that for a seven-storey non-ductile frame structure S_{de} is sufficient with respect to M_w and R when used for predicting the maximum

interstorey drift on the ground floor and the maximum floor acceleration at the roof level. Aslani [13] however notes that the sufficiency of S_{de} with respect to M_w will decrease as the influence of higher modes increases. This was illustrated by Luco and Cornell [7] who found S_{de} to be sufficient with respect to M_w for a 9-storey structure, but insufficient for a 20-storey structure. Aslani [13] illustrated that an IM equivalent to *PGA* was insufficient with respect to M_w for predicting maximum interstorey drifts, but is sufficient with respect to R. *PGA* was also found to be sufficient with respect to M_w and R for predicting maximum floor accelerations. Luco and Cornell [7] illustrate that S_{di} is sufficient with respect to both M_w and R when used for predicting peak interstorey drift ratio over all floors.

The majority of the above trends found by previous research were also observed for the seismic response analysis of the case study structure presented in this chapter. Figure 5-7 illustrates the sufficiency of the five different IMs with respect to M_w and R when predicting peak interstorey drifts and peak floor accelerations. Similar to the variation in efficiency of the different IMs over the height of the structure, there is also variation in sufficiency of the IMs for different EDPs. For example, S_{di} is found to be less sufficient with respect to M_w and R in the upper floors of the structure, where higher mode effects are more prominent (as S_{di} only captures inelastic period-lengthening effects). Both S_{de} and S_{di} were found to be less sufficient for predicting peak floor accelerations in the structure (having p-values less than 0.05 with respect to R in Figure 5-7d), which is controlled by higher mode vibrations. *PGA* is sufficient for both M_w, and R at predicting peak floor accelerations but less sufficient at predicting peak interstorey drifts in the central half (i.e. storeys 2-7) of the structure which is dominated by the first-mode response. The velocity-based IMs (*PGV* and *SI*) are generally found to be sufficient with respect to M_w and R for predicting peak interstorey drifts, and relatively sufficient at predicting peak floor accelerations (sufficient with respect to M_w, but marginally insufficient with respect to R). Thus it appears to be a clear trend that the better the efficiency of an IM at predicting a specific EDP, the better its sufficiency with respect to M_w and R.

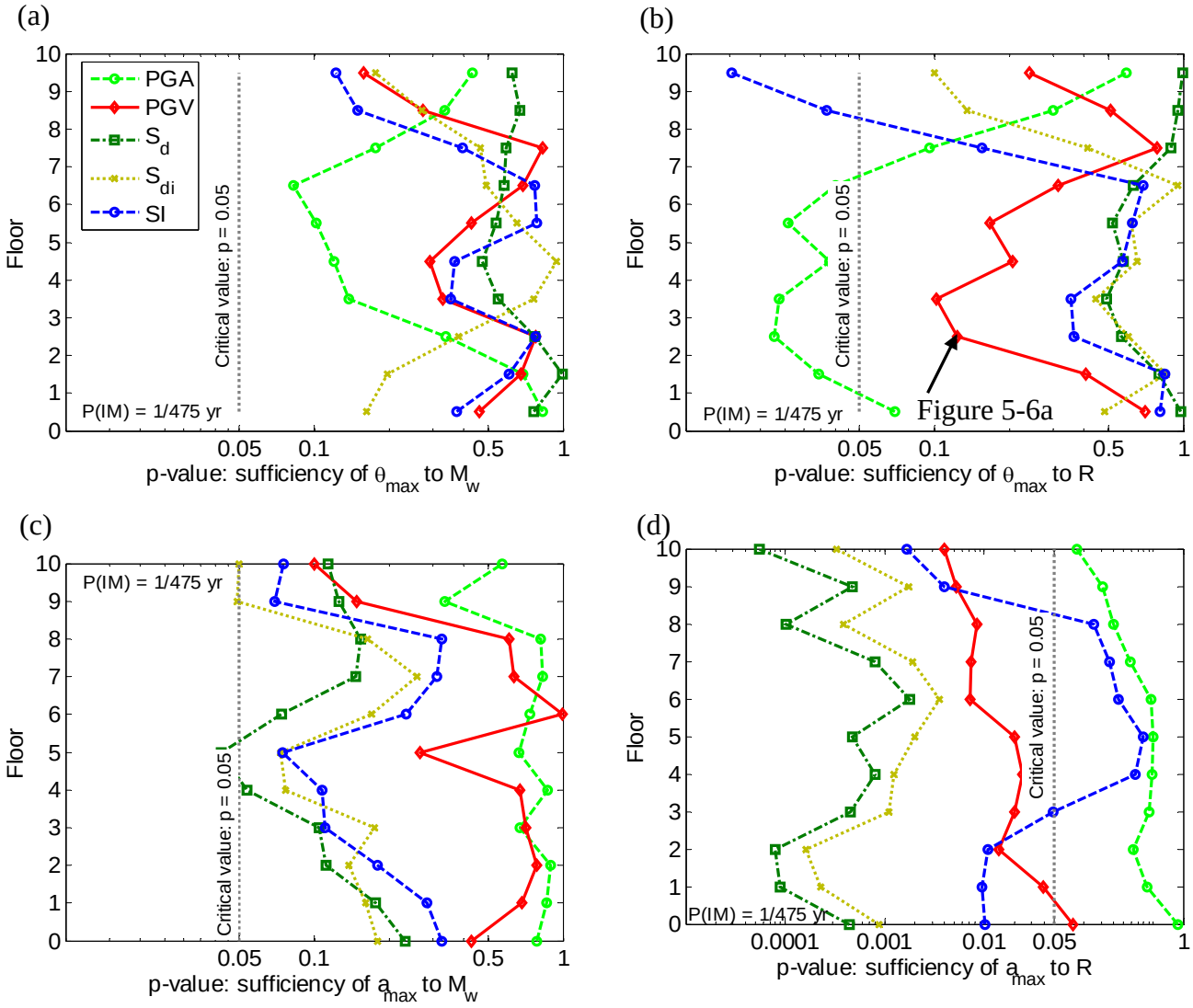


Figure 5-7: Sufficiency of the various IMs with respect to magnitude and distance when predicting peak interstorey drifts and peak floor accelerations.

To explain the relationship between efficiency and sufficiency described above assume that EDP is a deterministic function of multiple explanatory variables; $EDP = f(\boldsymbol{\Theta})$. Uncertainty in the EDP for a given IM occurs because the IM does not account for all of the explanatory variables which influence EDP. Now if a particular IM correlates strongly with a large majority of the explanatory variables, $\boldsymbol{\Theta}$, then given IM the uncertainty in EDP will be relatively small (i.e. IM will be efficient at predicting EDP). Because of the strong correlation of IM and many of the explanatory variables, then the response conditioned on IM will already implicitly account for many of the other explanatory variables and hence EDP conditioned on IM will be (practically speaking) independent (i.e. sufficient) with respect to these explanatory variables.

5.8.2 Sufficiency with respect to 'epsilon'

As previously noted, the parameter epsilon (ε) is defined as the number of standard deviations a specific ground motion parameter is above the mean predicted by a ground motion prediction equation. Epsilon is one of three parameters which can be obtained from seismic hazard deaggregation (the other two being M_w and R). In particular, because of the partial-correlated nature of spectral acceleration ordinates at different periods, Baker and Cornell [6] have shown that when ground motion records are scaled to a common $S_a(T)$ value (this includes S_{de} and PGA of the IMs examined here), ε can be used as a proxy for spectral shape, and therefore the extent to which higher-modes and period elongation will affect the seismic response for a given ground motion. Tothong and Luco [5] illustrate why S_{di} accounts for spectral shape at periods longer than that which the ground motions are scaled to (but not for spectral shape at higher mode periods). While Baker and Cornell [6] and Tothong and Luco [5] focus on the effect of epsilon when investigating peak interstorey drift ratios, the same logic can be directly applied to its effect on peak floor accelerations.

Figure 5-8a illustrates the sufficiency of the five IMs with respect to ε, when predicting peak interstorey drifts. Since when ground motion records are scaled to PGA, ε (being an indicator of spectral shape) will directly relate to the magnitude of the response spectra at longer periods, then PGA is insufficient with respect to ε for predicting interstorey drifts (which are mostly dominated by first mode vibration). Similarly, S_{de} and S_{di} are insufficient with respect to ε in predicting peak floor accelerations and peak interstorey drifts in upper floors which are dominated by higher mode vibration. PGV and SI are found to be sufficient to ε for predicting peak interstorey drifts, but insufficient for peak floor accelerations, as illustrated in Figure 5-8b.

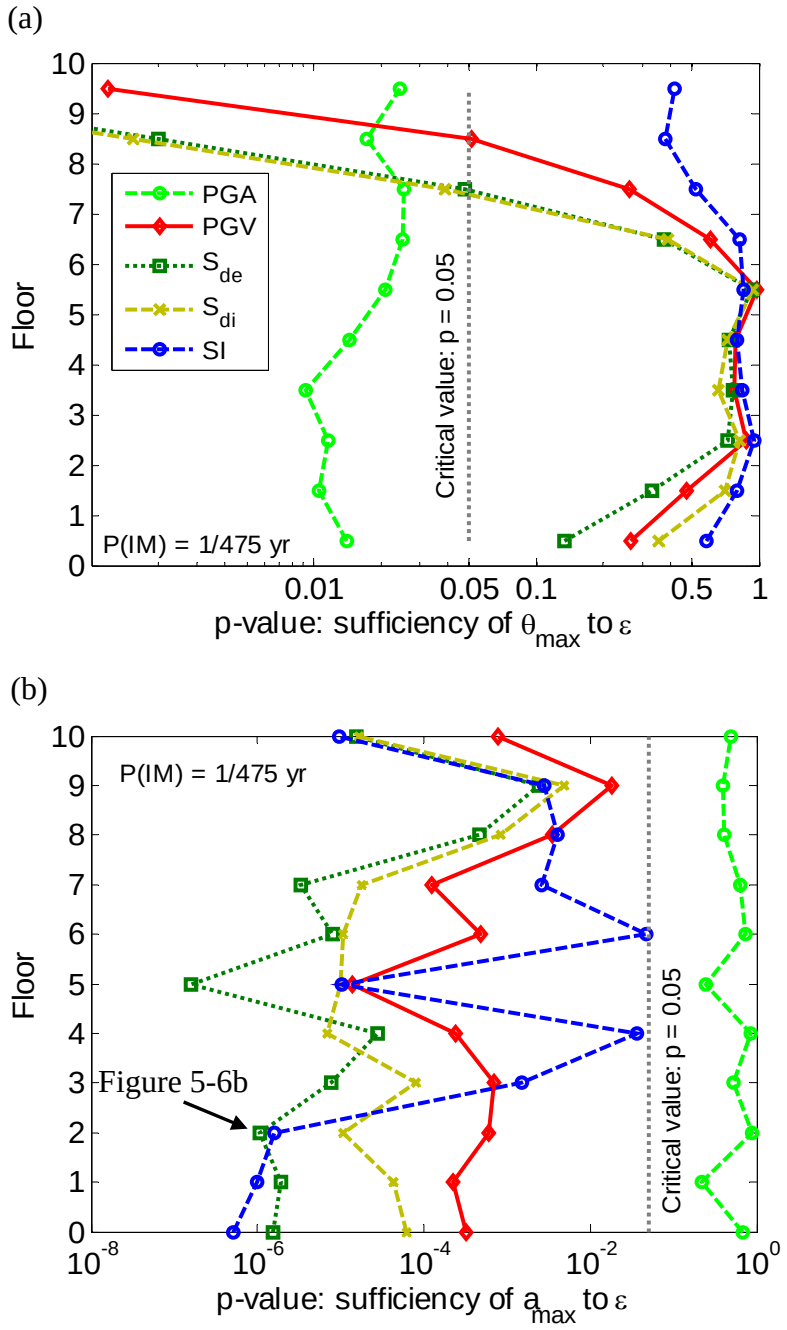


Figure 5-8: Sufficiency of the various IMs with respect to epsilon, ε.

5.8.3 Sufficiency with respect to scale factor (SF)

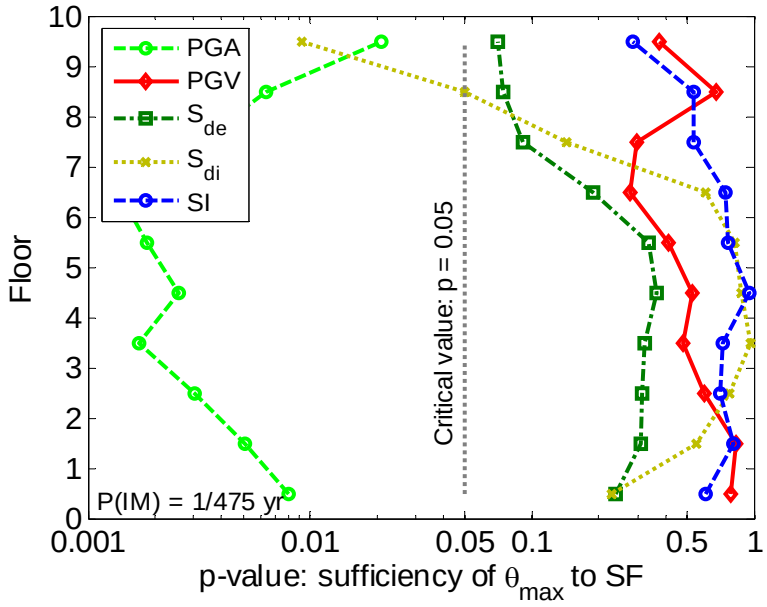
The use of linear scaling (i.e. scaling amplitude without modifying frequency content or duration) ground motion records to a specific IM level has been scrutinized for introducing bias compared with some 'true' response that would be obtained using un-scaled ground motions. Luco and Bazzurro [11] have illustrated that when using S_{de} as an IM, scale factors that are significantly different from 1.0 can introduce significant bias in seismic response analysis. Baker [38] has shown however that careful ground motion selection (similar to what

is used here) can significantly reduce such bias. Figure 5-9 illustrates the sufficiency of the five IMs with respect to SF for predicting peak interstorey drifts and peak floor accelerations. It can be seen that despite *PGA* for peak interstorey drifts, and *PGV* for peak floor accelerations, there is little dependence of the IMs on SF. As mentioned above, this is likely the result of ground motion selection based on deaggregation of the seismic hazard. Also, because of the ground motion selection employed here, the scale factors required to scale the ground motions are not significantly large, with means of 2.1 and 3.2 (averaged over all five IMs) and ranges of 0.6-5.7 and 0.81-9.1 for the 1/475 and 1/2475 exceedance probabilities, respectively.

5.8.4 Sufficiency with respect to site shear wave velocity, Vs_{30}.

With contemporary ground motion prediction equations [e.g., 20] giving soil site classification in terms of a 30-m averaged shear wave velocity, Vs_{30}, as opposed to a qualitative alphabet-based classification, it is possible to investigate the sufficiency of the various IMs with respect to Vs_{30}. Figure 5-10 illustrates the sufficiency of the IMs investigated with respect to Vs_{30} for both peak interstorey drifts and peak floor accelerations. In general all IMs are sufficient with respect to Vs_{30}. It is also worth noting that the considered range Vs_{30} = 300 – 800 m/s represents relatively stiff soils. Further studies are needed to investigate the sufficiency with respect to Vs_{30} for soft soil deposits (Vs_{30}<300 m/s) and significant nonlinearity in the soil response

(a)



(b)

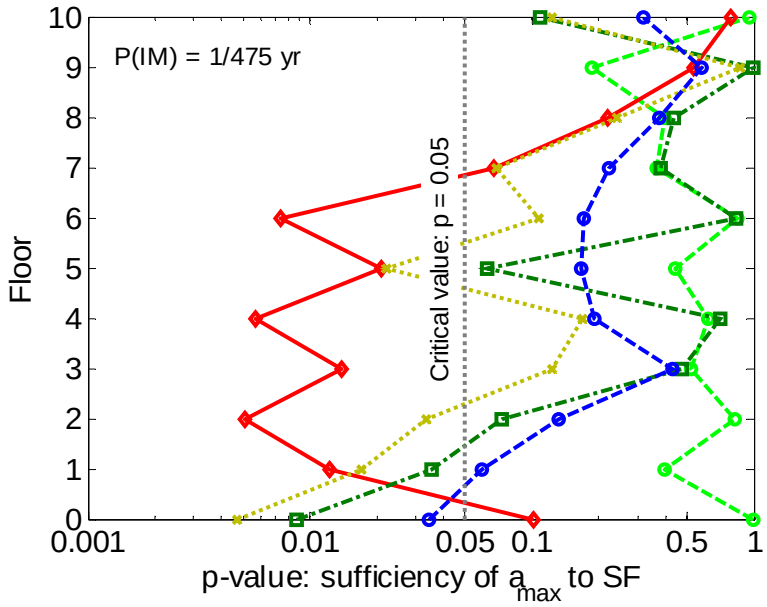


Figure 5-9: Sufficiency of the various IMs with respect to scale factor, *SF*.

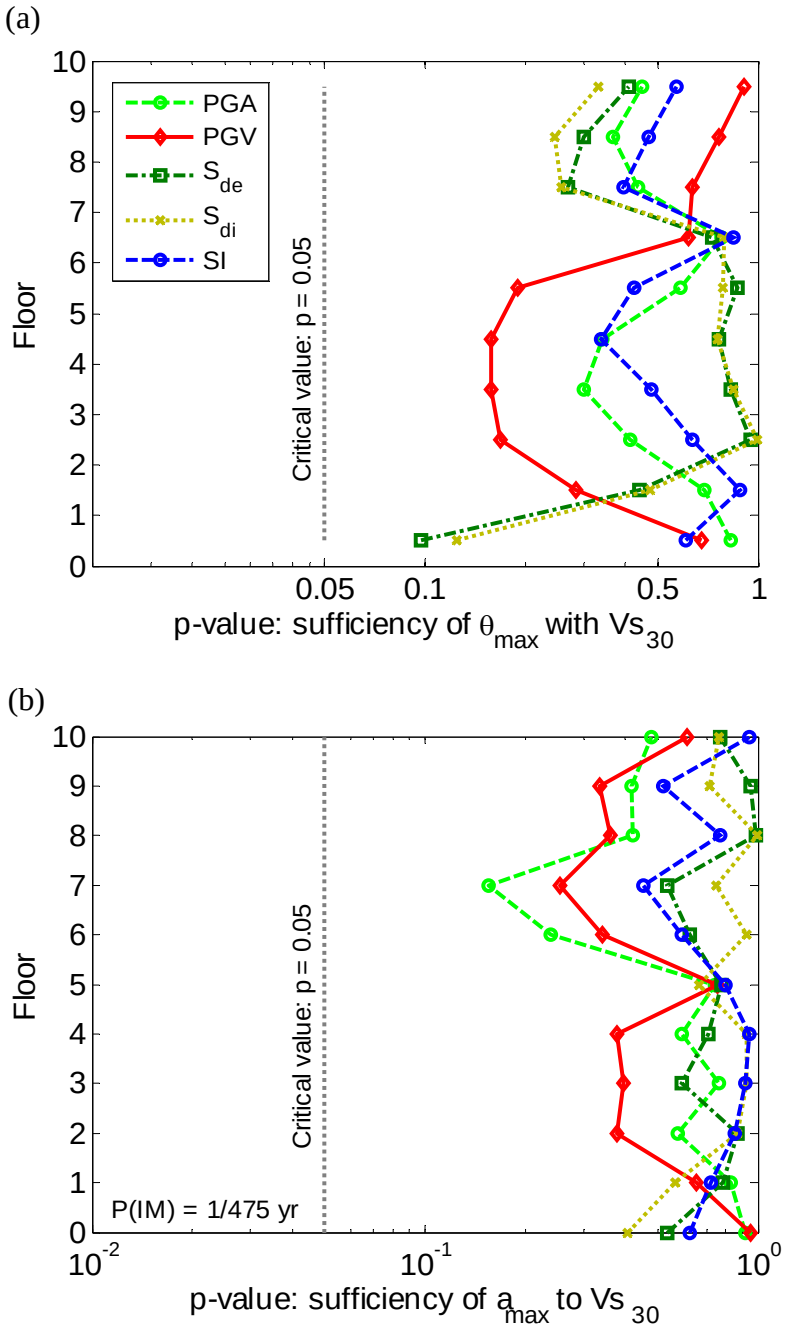


Figure 5-10: Sufficiency of the various IMs with respect to 30m-averaged site shear wave velocity, Vs_{30}.

5.8.5 'Correcting' seismic demand distributions

The aforementioned results have clearly illustrated that no single IM is sufficient with respect to M_w, R, ε , for all drift and floor acceleration demands in the case study structure. Therefore bias will be introduced in the computed response when the (M_w, R, ε) distribution of the ground motion suite is different than that of the hazard deaggregation when predicting EDPs using an insufficient IM. It was also noted that it may be difficult to obtain a desired

number of ground motion records which match the hazard deaggregation. In such cases, Bradley *et al.* [23] show that it is possible to regress on the results of the seismic response analyses to 'correct' the resulting distribution of EDP to reflect the M_w, R, ε distribution from hazard deaggregation.

An alternative option is to avoid the "IM approach" and simply develop a prediction equation for the vector of EDPs directly as a function of ground motion parameters (e.g. magnitude, distance) (e.g. [38]). While this approach will not suffer from some of the problems of the IM approach discussed in this manuscript it should be made clear that the development of such a prediction equation for a vector of EDPs is complex, requiring significantly more seismic response analyses, complex regression analyses, expertise in strong ground motion modelling and access to ground motion metadata and earthquake fault databases. The IM approach bypasses these difficulties by separating seismic hazard analysis and seismic response analysis. Both approaches have their respective pros and cons and will thus be useful in different situations.

5.9 Conclusions

Prediction of the seismic response of multi-degree-of-freedom structures is a complex task due to the spatially distributed seismic demands which are sensitive to different frequency contents of the imposed ground motion. This chapter investigated the efficacy of five different ground motion intensity measures (IMs): peak ground acceleration, *PGA*; peak ground velocity, *PGV*; elastic and inelastic spectral displacement, S_{de}, S_{di}; and spectrum intensity, *SI*, for which robust ground motion prediction equations are available. The concepts of predictability, efficiency, and sufficiency were investigated for each of the IMs when applied to the seismic response analysis of a 10-storey RC frame structure. It was illustrated that the efficiency (uncertainty in seismic response prediction) of an IM can be qualitatively determined based on the frequency range of the ground motion which controls the IM to that which controls the EDP being monitored. The predictability (uncertainty in ground motion prediction) of an IM was clearly shown to be an important factor in reducing the median response demand for ground motions scaled to an IM with a given probability of exceedance. Ground motion IMs which are predictable (namely *PGA*, *PGV* and *SI*), result in lower median seismic demands when ground motions are scaled to a specific probability of exceedance using a seismic hazard curve. The sufficiency of the IM (the dependence of the seismic response on parameters such as M_w, R, and ε) was found to be closely related to the efficiency

of the IM (the more efficient the IM, the higher the sufficiency), and it was observed that all IMs were insufficient with respect to at least one of M_w, R, or ε when predicting both peak interstorey drifts and peak floor accelerations. As a result of the insufficiency of an IM, careful ground motion selection (compatible with the hazard deaggregation) and/or appropriate response modification is therefore needed to reduce bias and dispersion.

5.10 Acknowledgements

Financial support of the first author from the New Zealand Tertiary Education Commission Bright Futures scheme is appreciated.

5.11 References

[1] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).

[2] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. *Report No. RMS-35, RMS Program*,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[3] Luco N. Probablistic Seismic Demand Analysis, Connection Fractures, and Near-Source Effects. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2002, 285.

[4] Kramer SL and Mitchell RA. Ground motion intensity measures for liquefaction hazard evaluation. *Earthquake Spectra* 2006; **22**(2): 413-438.

[5] Tothong P and Luco N. Probabilistic Seismic Demand Analysis Using Advanced Ground Motion Intensity Measures. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1837-1860.

[6] Baker JW and Cornell CA. A vector-valued ground motion intensity measure consisting of spectral acceleration and eplison. *Earthquake Engineering and Structural Dynamics* 2005; **34**(10): 1193-1217.

[7] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra* 2007; **23**(2): 357-392, DOI: 10.1193/1.2723158.

[8] Shome N, Cornell CA, Bazzurro P, and Carballo JE. Earthquakes, records, and nonlinear responses. *Earthquake Spectra* 1998; **14**(3): 469-500.

[9] Vamvatsikos D and Cornell CA. Developing efficient scalar and vector intensity measures for IDA capacity estimation by incorporating elastic spectral shape information. *Earthquake Engineering and Structural Dynamics* 2005; **34**(13): 1573-1600.

[10] Baker JW and Cornell CA. Spectral Shape, record selection and epsilon. *Earthquake Engineering and Structural Dynamics* 2006; **35**(9): 1077-1095.

[11] Luco N and Bazzurro P. Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses? *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1813-1835, DOI: 10.1002/eqe.695.

[12] Taghavi S and Miranda E. Probabilistic study of peak floor acceleration demands in linear structures, in *9th International Conference on Applications of Statistics and Probability in Civil Engineering*, San Francisco, CA, 2003, 8.

[13] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[14] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: from structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2008: (submitted).

[15] Bull DK and Brunsdon D. Examples of Concrete Structural Design to New Zealand Standards 3101. New Zealand, 1998.

[16] Standards New Zealand. NZS 3101 1995: Part 1: Concrete Structures Standard. Wellington, NZ, 1995.

[17] McKenna F, Fenves GL, and Scott MH. OpenSees: Open System for Earthquake Engineering Simulation.: Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA, 2004.

[18] Kramer SL. *Geotechnical Earthquake Engineering*. Prentice-Hall: Upper Saddle River, NJ., 1996; 653.

[19] Peng B, Dhakal R, Fenwick R, Carr A, and Bull D. Experimental investigation on the interaction of reinforced concrete frames with precast-prestressed concrete floor systems. *14th World Conference on Earthquake Engineering*, Beijing, China, 2008.

[20] Power M, Chiou B, Abrahamson NA, Bozorgnia Y, Shantz T, and Roblee C. An Overview of the NGA Project. *Earthquake Spectra* 2008; **24**(1): 3-21.

[21] Tothong P and Cornell CA. An Emperical Ground Motion Attenuation Equation for Inelastic Spectral Displacement. *Bulletin of the Seismological Society of America* 2006; **96**(6): 2146-2164.

[22] Housner GW. Spectrum intensities of strong-motion earthquakes, in *Symposium on earthquakes and blast effects on structures*, Los Angeles, CA, 1952.

[23] Bradley BA, Cubrinovski M, MacRae GA, and Dhakal RP. Ground motion prediction equation for Spectum intensity from spectral acceleration relationships. *Bulletin of the Seismological Society of America (in press)* 2008.

[24] Boore DM and Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s. *Earthquake Spectra* 2008; **24**(1): 99-138.

[25] Field EH, Jordan TH, and Cornell CA. OpenSHA: A Developing Community-Modelling Environment for Seismic Hazard Analysis. *Seismological Research Letters* 2003; **74**: 406-419.

[26] Boore DM, Watson-Lamprey J, and Abrahamson NA. Orientation-Independent Measures of Ground Motion. *Bulletin of the Seismological Society of America* 2006; **96**(4A): 1502-1511, DOI: 10.1785/0120050209.

[27] Goulet CA, Watson-Lamprey J, Baker JW, Haselton CB, and Luco N. Assessment of Ground Motion Selection and Modification (GMSM) methods for non-linear dynamic analyses of structures. *Geotechnical Earthquake Engineering and Soil Dynamics IV (GSP 181)*, Sacramento, CA, 2008.

[28] McGuire RK. Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bulletin of the Seismological Society of America* 1995; **85**(5): 1275-1284.

[29] Bazzurro P and Cornell CA. Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America* 1999; **89**(2): 501-520.

[30] Chiou B, Darragh R, Gregor N, and Silva WJ. NGA Project Strong-Motion Database. *Earthquake Spectra* 2008; **24**(1): 23-44.

[31] Baker JW and Cornell CA. Which Spectral Acceleration Are You Ysing? *Earthquake Spectra* 2006; **22**(2): 293-312.

[32] Bommer JJ and Abrahamson NA. Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates? *Bulletin of the Seismological Society of America* 2006; **96**(6): 1967-1977, DOI: 10.1785/0120060043.

[33] McGuire RK. *Seismic Hazard and Risk Analysis*. Earthquake Engineering Research Institute, 2004; 221.

[34] Tothong P. Probabilistic Seismic Demand Analysis using Advanced Ground Motion Intensity Measures, Attenuation Relationships and Near Fault effects. PhD thesis, Department of Civil and Environmental Engineering Stanford University, 2007,

[35] Beyer K and Bommer JJ. Selection and Scaling of Real Accelerograms for Bi-Directional Loading: A Review of Current Practice and Code Provisions. *Journal of Earthquake Engineering* 2007; **11**(1): 13-45, DOI: 10.1080/13632460701280013.

[36] Aslani H and Miranda E. Probability-based Seismic Response Analysis. *Engineering Structures* 2005; **27**(8): 1151-1163.

[37] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[38] Baker JW. Measuring Bias in Structural Response Caused by Ground Motion Scaling, in *8th Pacific Conference on Earthquake Engineering*, Singapore, 2007, 8.

6. Prediction of Spatially Distributed Seismic Demands in Specific Structures: Structural Response to Loss Estimation

Bradley BA, Dhakal RP, MacRae GA, Cubrinovski M. Prediction of spatially distributed seismic demands in structures: structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2009. (submitted).

6.1 Abstract

A companion chapter has investigated the effects of intensity measure (IM) selection in the prediction of spatially distributed response in a multi-degree-of-freedom structure. This chapter extends from structural response prediction to performance assessment metrics such as: probability of structural collapse; probability of exceeding a specified level of demand or direct repair cost; and the distribution of direct repair loss for a given level of ground motion. In addition, a method is proposed to account for the effect of varying seismological properties of ground motions on seismic demand that does not require different ground motion records to be used for each intensity level. Results illustrate that the conventional IM, spectral displacement at the first mode, $S_{de}(T_1)$, produces higher risk estimates than alternative velocity-based IM's, namely spectrum intensity, *SI*, and peak ground velocity, *PGV*, because of its high uncertainty in ground motion prediction and poor efficiency in predicting peak acceleration demands.

6.2 Introduction

Contemporary building-specific seismic performance and loss estimation methodologies [e.g. 1, 2, 3] use ground motion intensity measures (IM's) as the link between seismic hazard and structural response. Such methodologies also account for the spatially distributed location of components in the structure by using a vector of seismic demands (so-called engineering demand parameters, EDP's) to describe the response. It is therefore desirable to have an IM that can provide a precise (efficient) and unbiased (sufficient) prediction of the spatially distributed demands in structures. In addition, it is also important that the IM can itself be predicted with relatively small uncertainty (predictability).

Bradley *et al.* [4] investigated the efficiency, sufficiency, and predictability of common IM's when conducting seismic response analysis of a 10 storey RC frame structure. Bradley *et al.* illustrated that common ground motion IMs are not sufficient with respect to all of earthquake magnitude, M_w, source-to-site distance, R, and epsilon, ε, when predicting peak interstorey drift ratios and peak floor accelerations at spatially distributed locations in a structure, and thus bias will be introduced if the M_w, R, ε distribution of the ground motion suite used is different than that of seismic hazard deaggregation.

This chapter is intended to follow-on from Bradley *et al.* [4] with particular focus on: (1) 'correction' of the seismic demand distributions to account for insufficiency of the adopted IM; (2) determination of the annual probability of exceeding a specified level of demand (demand hazard); (3) determination of the distribution of direct economic loss for a given level of IM; (4) deaggregation of the expected direct economic loss for a given level of IM; and (5) determination of the annual probability of exceeding a specified level of direct economic loss (loss hazard). Explicit discussion is also given to the efficiency, sufficiency, and predictability of intensity measures in cases where results presented oppose current ideas as published in literature.

The structure used in the analyses presented herein is a 10 storey RC frame structure which is further described in Bradley *et al.* [4].

6.3 Hazard deaggregation and ground motion selection for a range of IM levels

Ground motion intensity measures (IM's) provide the link between the exceedance rate

of various levels of ground motion (from a seismic hazard curve) and structural response (from seismic response analysis). Consideration of an appropriate IM is discussed and investigated for the seismic response of a 10 storey RC frame in a companion chapter [4]. The five different IM's used here are: peak ground acceleration (*PGA*); peak ground velocity (*PGV*), elastic spectral displacement (S_{de}); inelastic spectral displacement (S_{di}); and spectrum intensity (*SI*). This selection was based on: (1) IMs which have been used by other researchers; and (2) IMs that have 'robust' ground motion prediction equations which can be used to determine ground motion hazard curves for the adopted IM. Bradley *et al.* [4, Figure 1] illustrate the hypothetical site and the seismic hazard curves obtained for the five different IMs.

Selection of ground motions for seismic response analysis should be based on those which are most likely to occur at the site in the future (which obviously is unknown). Deaggregation of the seismic hazard for a given level of ground motion, IM = *im*, can be used to determine the magnitude (M_w), distance (R), and epsilon (ε) statistics of the ground motion hazard, which can then be used as a 'target' for selecting ground motion records. Because different rupture scenarios (M_w, R, ε) have different probabilities of occurrence then deaggregation of the seismic hazard is dependent on the specific value of the IM considered. The seismic hazard deaggregation used herein gives the contribution to the total hazard of the various rupture scenarios for IM = *im*, i.e. $P\left(M_w, R, \varepsilon \middle| IM = im\right)$. As deaggregation of the ground motion hazard gives $P\left(M_w, R, \varepsilon \middle| IM > im\right)$ directly, then $P\left(M_w, R, \varepsilon \middle| IM = im\right)$ can be obtained by [5]:

$$P\left(M_w, R, \varepsilon \middle| IM = im_i\right) \approx \frac{1}{\Delta P\left(IM = im_i\right)}\left[P\left(M_w, R, \varepsilon \middle| IM > im_{i-1}\right)P\left(IM > im_{i-1}\right) - P\left(M_w, R, \varepsilon \middle| IM > im_{i+1}\right)P\left(IM > im_{i+1}\right)\right] \quad (6\text{-}1)$$

where $\Delta P\left(IM = im_i\right) = P\left(IM > im_{i-1}\right) - P\left(IM > im_{i+1}\right)$ and $im_{i-1} < im_i < im_{i+1}$.

The fact that the results of seismic hazard deaggregation are a function of the IM level (as well as the IM adopted), indicates that if a wide range of IM levels is considered, then the deaggregation statistics may change significantly. Thus, different ground motion record suites should be used for different IM levels. Such an approach was adopted by, for example, Goulet *et al.* [3] who used seven different ground motions suites for seven IM levels from 0.1-0.82g $S_a(T=1.0\text{s})$. The need to use different ground motion suites at different IM levels to reduce bias is conditional on the insufficiency of the adopted IM to a specific parameter (i.e. M_w, R, ε) from the hazard deaggregation. If an IM is sufficient for all M_w, R, ε then the

seismic response of some structure to ground motions scaled to IM is conditionally independent of M_w, R, ε. However, Bradley *et al.* [4] illustrate that common IMs used in seismic response analysis do not exhibit sufficiency with respect to all of M_w, R, ε for predicting peak interstorey drift ratios and peak floor accelerations.

Figures 6-1a-6-1c illustrate the mean and mean ± one standard deviation values of M_w, R, and ε obtained from hazard deaggregation for the site as a function of PGA annual exceedance probability. It is seen that as the probability of exceedance reduces, resulting in larger ground motions, both the mean magnitude and epsilon values increase, while the distance decreases. Also shown in Figures 6-1a-6-1c are the mean and mean ± one standard deviation values of M_w, R, and ε of the ground motion suite used by Bradley *et al.* [4], which was also adopted in this study for all intensity levels. It is clear that the properties of the ground motion suite can be quite different than the hazard deaggregation as the exceedance probability varies. Figures 6-1d-6-1f illustrate the variation in the mean values of M_w, R, and ε for the five different IMs as a function of the probability of exceedance of the IM. While the same trends as mentioned above for PGA apply to all of the IMs, there is clearly variation in the mean values for a given probability of exceedance. This can be explained with reference to the usual scenario in which small magnitude events at close distances dominate a uniform hazard spectra (UHS) at short vibration periods (e.g. *PGA*), while larger magnitude events at farther distances tend to dominate the UHS at longer vibration periods (e.g. S_{de} and S_{di}). In this study, the seismic response of the case study structure is investigated over a range of IM levels. Rather than using different ground motion suites for different IMs and IM levels, a single ground motion suite is used, and the distribution of the seismic demand obtained is 'corrected' to account for the difference between the deaggregation and ground motion suite M_w, R, and ε statistics. Bradley *et al.* [4] illustrate that even for a given IM level it maybe difficult to obtain a ground motion suite where the M_w, R, and ε statistics closely match some 'target' statistics from deaggregation. In such cases, 'correction' of the seismic demand distribution due to IM insufficiency may also be necessary (particularly for loss estimation, where typically both interstorey drifts and floor accelerations are estimated). As will be shown in the following section, the effect of the 'correction' depends on the insufficiency of an IM with respect to (one or more of) M_w, R, and ε, and the difference between the deaggregation and ground motion suite statistics. Hence, with reference to Figures 6-1a-6-1c, in this study, the corrections will be most significant at low probabilities of exceedance where the largest difference between hazard deaggregation and ground motion properties is observed.

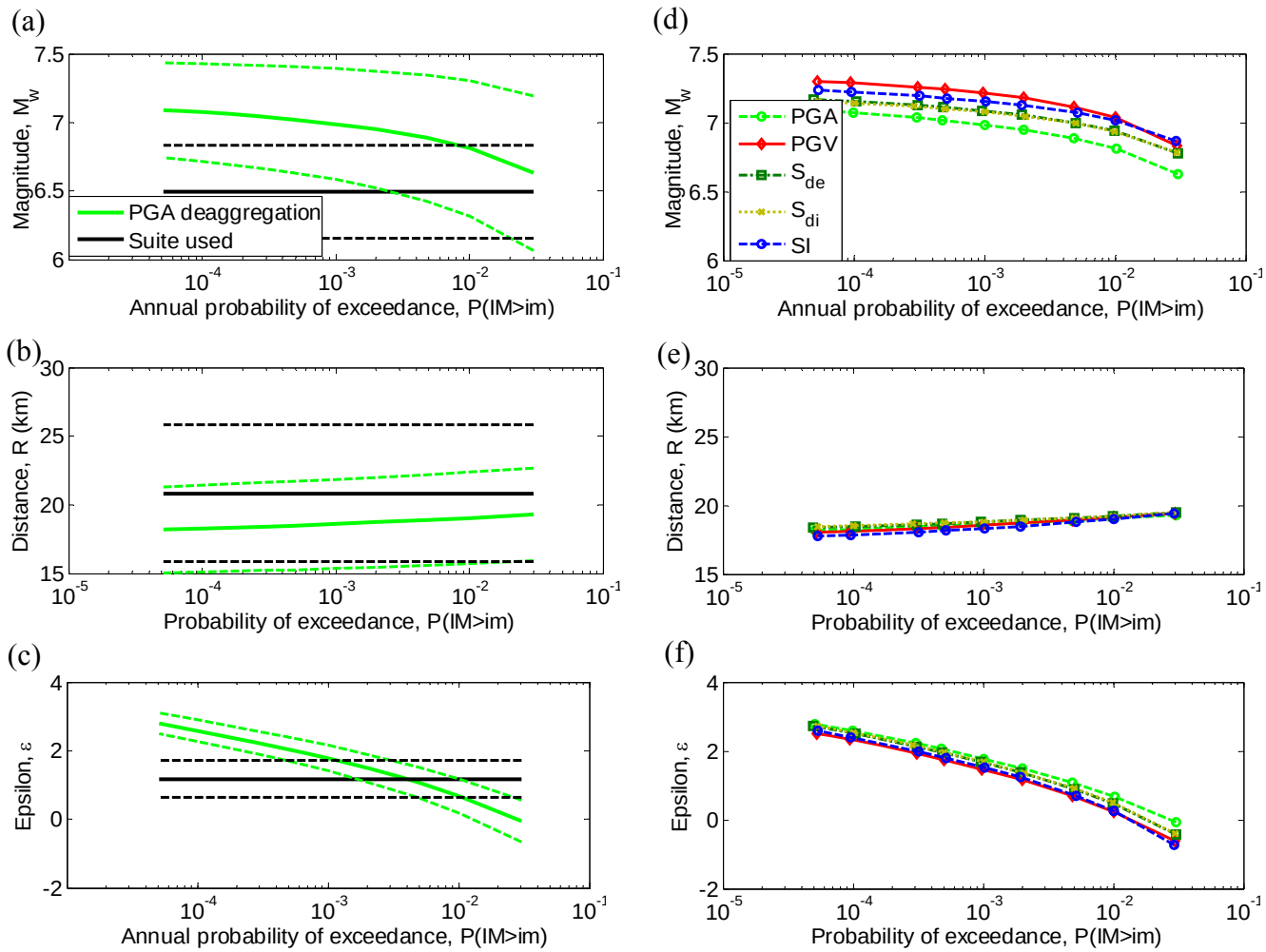


Figure 6-1: Comparison of the mean (solid line) and ± one standard deviation (dashed line) of: (a) magnitude; (b) source-to-site distance; and (c) epsilon values from hazard deaggregation of PGA with the ground motion suite used. Variation of the mean: (d) magnitude; (e) distance; and (f) epsilon with annual exceedance probability for the different IMs.

6.4 Correcting seismic demand distributions

Bradley *et al.* [4] illustrated that none of the IMs examined in this chapter are sufficient with respect to all of M_w, R, ε for all peak interstorey drifts and peak floor accelerations monitored in the case-study structure. The advantage of having a sufficient IM with respect to some parameter x, is that the distribution of demands obtained from the finite number of seismic response analyses, would not be (practically speaking) dependent on the distribution of x from the suite of ground motions used (i.e. $f(EDP \mid IM) \cong f(EDP \mid IM, x)$). However, as no IMs were entirely sufficient with respect to M_w, R, ε then in certain instances the properties of the ground motion suite used will affect the results and therefore careful ground motion selection is pertinent. As noted by Bradley *et al.* [4], despite the large database of ground

motion records available for time-history analysis it may not be possible to obtain a specified number of ground motions which accurately match the distribution of M_w, R, ε obtained from hazard deaggregation for the specified level of IM (particularly as $M_w$and ε increase and R reduces). In such cases it is possible to use the correlation between the observed EDPs and some parameter x to 'correct' the distribution of $f(EDP \mid IM)$. The theory behind such a procedure is outlined in the following paragraphs.

6.4.1 Theory of the demand correction procedure

From the result of multiple regression it is possible to obtain the mean of lnEDP as a function of M_w, R, ε (i.e. $\mu_{\ln EDP|IM,M_W,R,\varepsilon}$) and then integrate over the distribution of M_w, R, ε (i.e. $f(M_W,R,\varepsilon)$) from hazard deaggregation to obtain an unbiased estimate of the 'true' mean of the *lnEDP|IM* distribution as given in Equation (6-2):

$$\hat{\mu}_{\ln EDP|IM} = \iiint_{M_W,R,\varepsilon} \mu_{\ln EDP|IM,M_W,R,\varepsilon} f(M_W,R,\varepsilon) dM_W dR d\varepsilon \tag{6-2}$$

If it is however known that for the given EDP, IM is sufficient with respect to (for example) M_w and R, then $\mu_{\ln EDP|IM,M_W,R,\varepsilon} \cong \mu_{\ln EDP|IM,\varepsilon}$, and Equation (6-2) simplifies to:

$$\hat{\mu}_{\ln EDP|IM} \cong \int_{\varepsilon} \mu_{\ln EDP|IM,\varepsilon} f(\varepsilon) d\varepsilon \tag{6-3}$$

Similarly, the standard deviation of the *lnEDP|IM* distribution can be obtained (also assuming sufficiency with respect to M_w and R):

$$\hat{\sigma}^2_{\ln EDP|IM} \cong \int_{\varepsilon} \left(\mu^2_{\ln EDP|IM,\varepsilon}\right) f(\varepsilon) d\varepsilon - \left(\hat{\mu}_{\ln EDP|IM}\right)^2 \tag{6-4}$$

Furthermore if as a first-order approximation, $\mu_{\ln EDP|IM,\varepsilon}$ is assumed to be a linear function of ε and $f(\varepsilon)$ to have a normal distribution then Equations (6-3) and (6-4) become:

$$\hat{\mu}_{\ln EDP|IM} \approx \mu_{\ln EDP|IM,\varepsilon}(\mu_\varepsilon) \tag{6-5}$$

$$\hat{\sigma}^2_{\ln EDP|IM} \approx \sigma^2_{\ln EDP|IM,\varepsilon} + (b_\varepsilon \sigma_\varepsilon)^2 \tag{6-6}$$

where μ_ε and σ_ε are the mean and standard deviation of ε obtained from hazard deaggregation, and b_ε is the coefficient in the linear regression $E[\ln EDP \mid IM] = a + b_\varepsilon \varepsilon$. Equations (6-2)-(6-6) form the theoretical basis for the 'correction' method proposed by Haselton [6] for correcting the collapse capacity of structures due to insufficiency with respect to ε. As noted by Haselton [6], Equations (6-5) and (6-6) can: (i) correct bias introduced if the

distribution of the ground motion suite is significantly different from that of the hazard deaggregation; and (ii) potentially reduce dispersion (i.e. improve efficiency).

The following paragraphs generalise the result obtained above to the case of multiple insufficiency. When an IM is found to be insufficient with respect to multiple parameters then the multiple linear regression is of the form:

$$E\left[\ln EDP \middle| IM, \mathbf{X}\right] = a + \mathbf{b}_{\mathbf{X}} \bullet \mathbf{X} \tag{6-7}$$

where $\mathbf{X} = \left(M_W, R, \varepsilon\right)$ is a vector of the parameters for which the IM is insufficient; $\mathbf{b}_{\mathbf{X}} = \left(b_{M_W}, b_R, b_\varepsilon\right)$; and "$\bullet$" is the vector 'dot' product. It is trivial to generalise Equations (6-5) and (6-6) in this case to:

$$\hat{\mu}_{\ln EDP|IM} \approx \mu_{\ln EDP|IM,\mathbf{X}}\left(\boldsymbol{\mu}_X\right) \tag{6-8}$$

$$\hat{\sigma}^2_{\ln EDP|IM} \approx \sigma^2_{\ln EDP|IM,\mathbf{X}} + \boldsymbol{b}_X{}^{\boldsymbol{T}} \boldsymbol{\Sigma}_X \boldsymbol{b}_X \tag{6-9}$$

where $\boldsymbol{\Sigma}_X$ is the covariance matrix of the parameters in $\boldsymbol{X}$ obtained from hazard deaggregation; and "$^{\boldsymbol{T}}$" is the vector transpose. Thus, in the case of multiple insufficiency regression is performed over a vector of components, and knowledge of the mean and covariance matrix of the hazard deaggregation is necessary. Figure 6-2 illustrates the (marginal) hazard deaggregation of *PGA* at the 1/475 exceedance probability with respect to M_w and R (Figure 6-2a), and M_w and ε (Figure 6-2b). It can be seen from Figure 6-2a that there exists little correlation between M_w and R, with a correlation coefficient of $\rho_{M,R} = 0.09$. Figure 6-2b illustrates that M_w and ε however exhibit a strong negative correlation, with $\rho_{M,\varepsilon} = -0.89$. The reason for this significant negative correlation between M_w and ε is that they have the same effect on observed ground motions at a particular site (and thus must be negatively correlated to produce a ground motion with $IM = im$). Large M_w ruptures cause large ground motions which occur infrequently (compared to smaller M_w ruptures), while small ε values (i.e. around $\varepsilon = 0$) cause smaller ground motions which occur more frequently (compared to, for example $\varepsilon = 2$). Figure 6-2c illustrates the correlations between M_w, R, ε as a function of the exceedance probability of *PGA* for the site considered. While the correlations do vary with exceedance probability the general trends discussed above remain unchanged (similar trends were observed for the other IMs considered in this study). As the correlations are significant, then they should not be neglected when computing Equations (6-9) and (6-11) (which would lead to a significant underestimation of the 'corrected' dispersion of the ln*EDP*|*IM* distribution for the vector of EDP's).

In the case of loss estimation it is desired to not only know the marginal distributions of

all of the EDP's, but also to have knowledge of the joint distribution in the form of the correlation (or covariance) matrix of the EDPs. If the above equations are used to 'correct' the distribution for a single EDP, it is also necessary to use similar corrections for the correlation (or covariance) between the different EDPs. The covariance between $lnEDP_i$ and $lnEDP_j$ based on the multivariate regression equation (Equation (6-7)) is given by:

$$Cov\left[\ln EDP_i, \ln EDP_j \middle| IM\right] = Cov\left[\left(a_i + \mathbf{b}_{\mathbf{X}i} \bullet \mathbf{X} + \eta_i\right), \left(a_j + \mathbf{b}_{\mathbf{X}j} \bullet \mathbf{X} + \eta_j\right)\right] \tag{6-10}$$

where η_k are the regression residuals for $lnEDP_k$, which have a normal distribution with zero mean and standard deviation, $\sigma^2_{\ln EDP_k|IM,\mathbf{X}}$. Equation (6-10) can be expanded (using properties of the expectation operator, [7]) and simplified to obtain the corrected covariance between $lnEDP_i$ and $lnEDP_j$:

$$\hat{\sigma}_{\ln EDP_i, \ln EDP_j|IM} \approx \sigma_{\ln EDP_i, \ln EDP_j|IM,\mathbf{X}} + \boldsymbol{b}_{X_i}^{T} \boldsymbol{\Sigma}_X \boldsymbol{b}_{X_j} \tag{6-11}$$

where $\sigma_{\ln EDP_i, \ln EDP_j|IM,\mathbf{X}}$ is the covariance between the regression residuals η_i and η_j. In the case of $i = j$, as one would expect, Equation (6-11) simplifies to Equation (6-9).

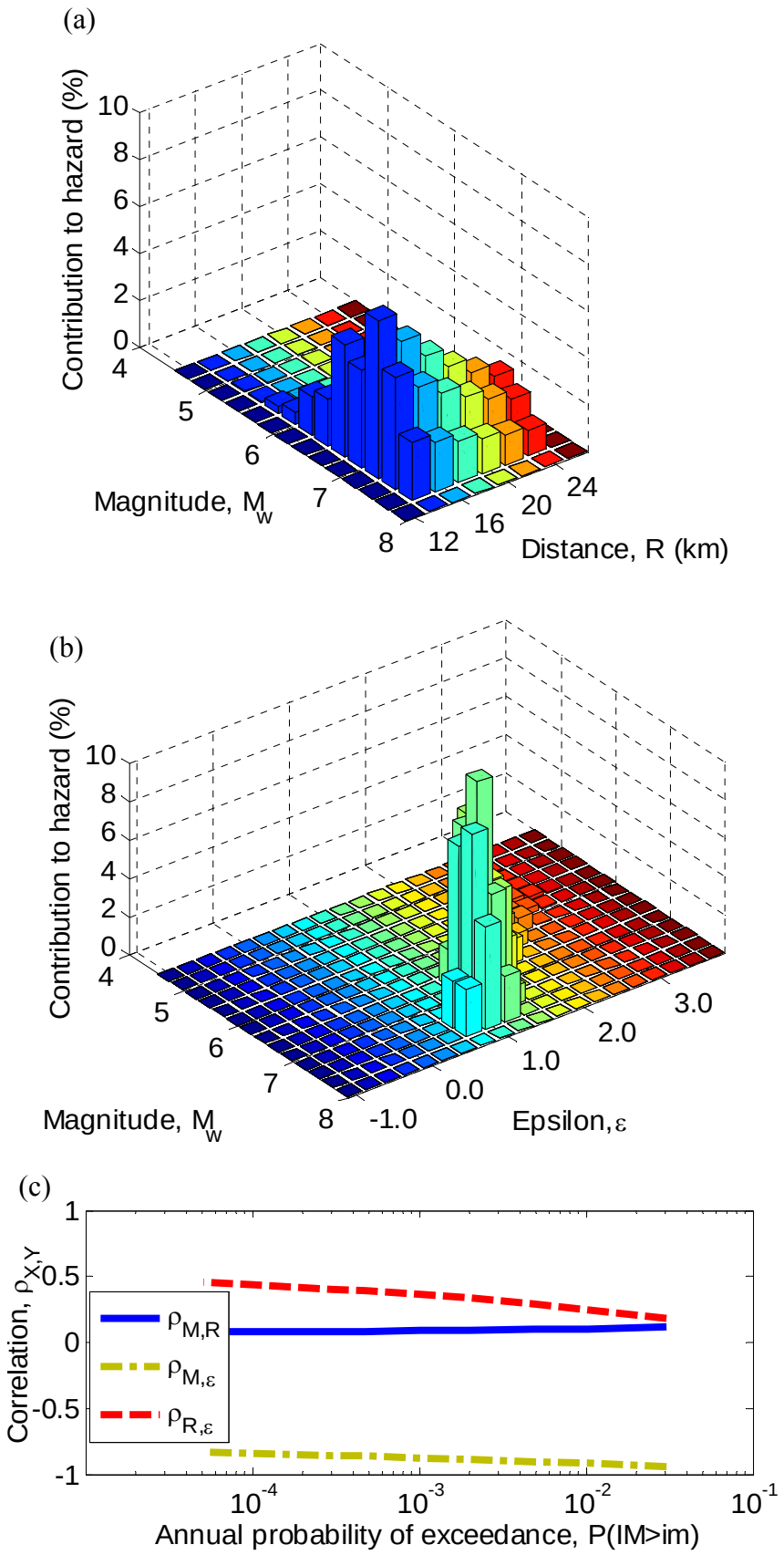


Figure 6-2: Illustration of the correlation within hazard deaggregation at the 1/475 exceedance probability for *PGA*: (a) magnitude and distance ($\rho_{M,R|IM}$ = 0.09); (b) magnitude and epsilon ($\rho_{M,\varepsilon|IM}$ = -0.89); and (c) variation in the deaggregation correlations as a function of *PGA* annual exceedance probability.

Collapse probability may also be an important decision variable for performance and loss estimation [6], and it is therefore also necessary to correct the probability of collapse. As previously mentioned, Haselton [6] illustrated that collapse capacity estimated using S_{de} is insufficient with respect to ε (and assumed sufficiency with respect to M_w, R). Haselton used an iterative procedure to determine the largest intensity of a specific ground motion at which the structure did not collapse, yielding 'collapse capacity' data which is continuous and can be 'corrected' using the aforementioned method. In the case of loss estimation, it is more desirable to use the seismic response analyses conducted with ground motions of IM = *im* (which are used to get the distribution of loss, $f(L \mid IM)$), to directly estimate the collapse probability. This however leads to binary data of collapse or no-collapse for each ground motion record which requires minor modifications to the above correction procedure. Logistic regression [8] is used here to regress on the binary collapse data. The cumulative distribution of the logistic random variable (the collapse probability) is given by:

$$P_{C|IM,\boldsymbol{X}}(\boldsymbol{X}) = \frac{\exp(a + \boldsymbol{b}_{\boldsymbol{X}} \bullet \boldsymbol{X})}{1 + \exp(a + \boldsymbol{b}_{\boldsymbol{X}} \bullet \boldsymbol{X})} \tag{6-12}$$

where $\boldsymbol{X}$ and $\boldsymbol{b}_{\boldsymbol{X}}$ have their same meanings as defined in Equation (6-7). The regression coefficients (i.e. $\boldsymbol{b}_{\boldsymbol{X}}$) are obtained by (numerically) maximising the likelihood function for the logistic density [8]. Once the regression coefficients have been determined then the 'corrected' probability of collapse, $\hat{P}_{C|IM}$ can be obtained from $\hat{P}_{C|IM} = P_{C|IM,\boldsymbol{X}}(\boldsymbol{\mu}_{\boldsymbol{X}})$.

As with the multiple linear regressions for correcting the demand distributions, significance tests should be performed on $\boldsymbol{b}_{\boldsymbol{X}}$ to confirm that there are statistically significant trends in the data. This particularly applies in the case where only a very small proportion of the ground motions cause structural collapse, which can produce statistically unstable regressions. If one or more variables in $\boldsymbol{X}$ are found to be statistically insignificant then the regression should be re-performed following removal of these variables.

While Equations (6-8), (6-9), (6-11), and (6-12) offer a method to correct the distribution of structural response, and collapse probability for an insufficient IM, the procedure to do so becomes complex in the case of insufficiency with respect to more than one parameter. In particular, knowledge of the correlation matrix between M_w, R, ε from the hazard deaggregation may not be available. In addition, for insufficiency with respect to multiple parameters the so-called *curse of dimensionality* [9] will mean that large suites of ground motions need to be used in order to obtain statistically stable 'corrected' results, as noted in a similar context by Baker [10]. Therefore, if a small suite of ground motion records

is to be used (such as that proposed in current code guidelines), ground motions should be selected to match the deaggregation of the seismic hazard. If 'exact' ground motion selection to match seismic hazard deaggregation is not possible, the above procedure can be used. Furthermore, adequate ground motion selection will likely mean that the effect of the 'corrections' is not overly excessive and may be neglected if such a level of accuracy is not required.

For the hypothetical scenario considered in this study (i.e. a single line source) the deaggregation at a given IM level is relatively simple. In reality however, it is common for the total seismic hazard at a site to have significant contributions from multiple sources which may mean that the deaggregated M_w, R, ε distribution is far from a multi-variate normal distribution, which has been assumed in the above correction procedure. In such cases, Equations (6-2) and (6-4), can be solved by summation over the deaggregation probability mass function (PMF) which is the typical output of seismic hazard deaggregation (e.g. Figures 6-2a and 6-2b).

6.4.2 Corrections for demand distribution and collapse probability

Figures 6-3a and 6-3b illustrate the regression of the maximum ground floor acceleration with M_w and R (a 3-dimensional regression with ε was performed, but can obviously not be plotted). Simply from inspection it can be seen that S_{de} is notably less sufficient to M_w and R compared to PGA. Table 6-1 illustrates the numerical values of the regression shown in Figures 6-3a and 6-3b. It can be seen that using PGA to predict the maximum ground floor acceleration

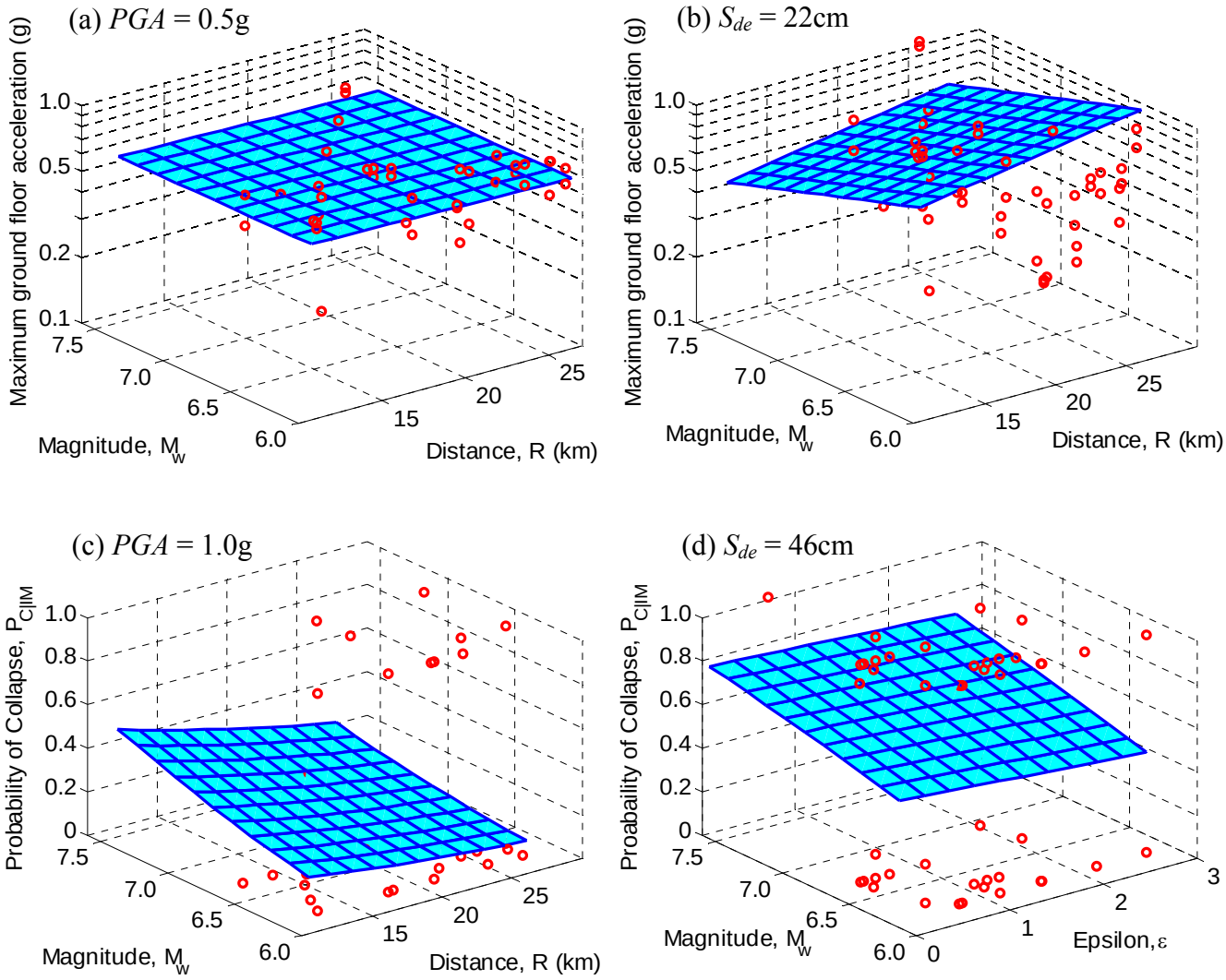


Figure 6-3: Illustration of: multivariate regression used to 'correct' the distribution of the demand given intensity for the maximum ground floor acceleration at the 1/475 probability of exceedance using: (a) *PGA*; and (b) S_{de}. Use of multivariate logistic regression to 'correct' the probability of structural collapse at the 10^{-4} exceedance probability using: (c) *PGA*; and (d) S_{de}.

($a_{max,1}$) the dependence on M_w, R and ε is statistically insignificant based on the F-test [8]. On the other hand, when S_{de} is used to predict $a_{max,1}$, R and ε are significant and the corrected and uncorrected values are exp(-0.569) = 0.57 g and exp(-0.527) = 0.59 g, respectively. It is noted that when viewing Table 6-1, the statistical significance of the 'b' values is dependent on both their magnitude as well as their uncertainty (due to the scatter in the data).

Figures 6-3c and 6-3d illustrate the logistic regression of the collapse probability ($P_{C|IM}$) on M_w and R, with numerical values given in Table 6-1. Figure 6-3d illustrates that there is some dependence of collapse capacity on M_w and ε when scaling ground motions based on S_{de}. Figure 6-3c also illustrates that there is a dependence on M_w, R when using PGA-scaling.

However, Table 6-1 illustrates that the R dependence is statistically insignificant (this is more easily examined by looking at the one-dimensional regressions against each of M_w, R and ε).

The dependencies observed in Figure 6-3 are somewhat intuitive, with higher magnitude ground motions having richer low frequency content and weaker (relatively speaking) high frequency content compared with smaller magnitude ground motions. The trends regarding ε are explained in detail by Baker and Cornell [11].

Results discussed in the following sections are all based on the use of the correction procedure (including statistical testing of significance) and hence for brevity the term 'corrected' is omitted.

Table 6-1: Illustration of correction procedure applied to the cases of Figure 6-3.

EDP	IM	a	$\mathbf{b_X}=(b_M, b_R, b_\varepsilon)$	$\boldsymbol{\mu_X}=(\mu_M, \mu_R, \mu_\varepsilon)$ (Figure 6-1)	$\hat{\mu}_{\ln EDP\|IM}$ (Eq (6-8)) or $\hat{P}_{C\|IM}$ (Eq (6-12)) corrected	$\mu_{\ln EDP\|IM}$ or $P_{C\|IM}$ uncorrected
$a_{max,1}$	PGA	-0.56	(-0.02, -0.001, 0.01)[2]	(6.95, 18.7, 1.49)	-0.686 (-0.685)[3]	-0.686
$a_{max,1}$	S_{de}	0.54	(-0.015[1],-0.03,-0.53)	(7.06,18.9,1.38)	-0.569	-0.527
$P_{C\|IM}$	PGA	0.32	(0.20, -0.08[1], -1.18)	(7.07, 18.3, 2.59)	0.26	0.21
$P_{C\|IM}$	S_{de}	-3.77	(0.69, -0.015, -0.24)	(7.15,18.5,2.49)	0.56	0.64

[1]coefficient is statistically insignificant based on t-test at α = 5% significance level
[2]all $\boldsymbol{b_X}$ coefficients are statistically insignificant based on F-test at α = 5% significance level
[3]values in brackets are those estimated using the (insignificant) regression equation

6.5 Collapse and demand hazards

Using the suite of ground motions adopted for this study, seismic response analyses were conducted by scaling the ground motion records to IM values that had annual exceedance probabilities ranging from 0.03 to $5\text{x}10^{-5}$. Figure 6-4 illustrates the mean scale factor of the ground motion suite (i.e. the average of the scale factors to scale each record to $IM = im$) which was required to scale the ground motions to IM values corresponding to the nine different exceedance probabilities. As one would expect, reducing the exceedance probability causes an increase in ground motion intensity and therefore an increase in the mean scale factor. It is also evident that for a given exceedance probability the mean scale factor of the ground motion suite is not the same when different IMs are used. Furthermore, the difference between the scale factors using different IMs (for a given exceedance probability) increases as the exceedance probability reduces. For example, at $P_{IM>im} = 0.03$ the ratio between the mean scale factors using S_{de} and SI is 0.77/0.66 = 1.17, while at $P_{IM>im} = 5\text{x}10^{-5}$ the ratio is 6.2/4.1 = 1.51. It should be clear from results presented in Bradley

et al. [4, Table 3] that the difference between the mean scale factor for the different IMs is strongly related to the predictability of the IM (i.e. the uncertainty in the ground motion prediction equation, GMPE), since a large uncertainty in the GMPE will significantly increase the hazard at low $P_{IM>im}$.

6.5.1 Probability of structural collapse

Within the PEER PBEE framework, performance with respect to structural collapse can be computed by combining the probability of collapse for a given IM, $P(C|IM = im)$, with the ground motion hazard curve, $P(IM > im)$ to obtain the probability of structural collapse, P_C [12]:

$$P_C = \int_{IM} P(C|IM = im)\left|\frac{P(IM > im)}{dIM}\right| dIM \tag{6-13}$$

$P(C|IM = im)$ is initially estimated from the proportion of ground motion records which cause structural collapse when scaled to IM = *im* and then corrected using Equation (6-12). As $P(C|IM = im)$ is generally assumed to have a lognormal distribution [12], then the values of $P(C|IM = im)$ found at the nine IM levels were used to determine the parameters of this distribution using a generalised linear model (GLM) with a probit link [8]. A GLM is used as it allows for non-constant variance as opposed to conventional Gaussian regression. Figure 6-5 illustrates the cumulative probability of structural collapse as a function of the IM exceedance probability. Important parameters of the lognormal distribution of collapse probability are given in Table 6-2. Firstly, Table 6-2 illustrates that the dispersion, $\sigma_{\ln IM|C}$, for the collapse fragility is highly correlated to the efficiency of the prediction of the maximum interstorey drifts over all floors of the structure [4], with S_{de} and S_{di} being the most efficient, followed by *SI*, *PGV* and lastly *PGA*. In particular, the large dispersion in the collapse fragility using *PGA* was not able to be reduced significantly using the 'correction' procedure as the logistic regression was insignificant when a small proportion of collapses occurred. Secondly, the collapse probability obviously depends on the predictability of the IM. For example at $P_{IM>im} = 5\text{x}10^{-5}$, using *SI* gives a collapse probability of approximately 0.34 compared to approximately 0.72 using S_{de}. Figure 6-4 indicates that for $P_{IM>im} = 5\text{x}10^{-5}$ the mean scale factor using *SI* is 4.1, while it is 6.2 using S_{de}. Using Figure 6-4 in an inverse manner gives a scale factor using S_{de} of approximately 4.1 when $P_{IM>im} = 3\text{x}10^{-4}$, an

exceedance probability at which the collapse probability is approximately 0.31 using S_{de} (Figure 6-5). Hence, for a scale factor of approximately 4.1, *SI* and S_{de} give collapse probabilities of 0.34 and 0.31, respectively. Thus, with the exception of *PGA*, the significant difference between the collapse distributions is primarily due to the IM predictability.

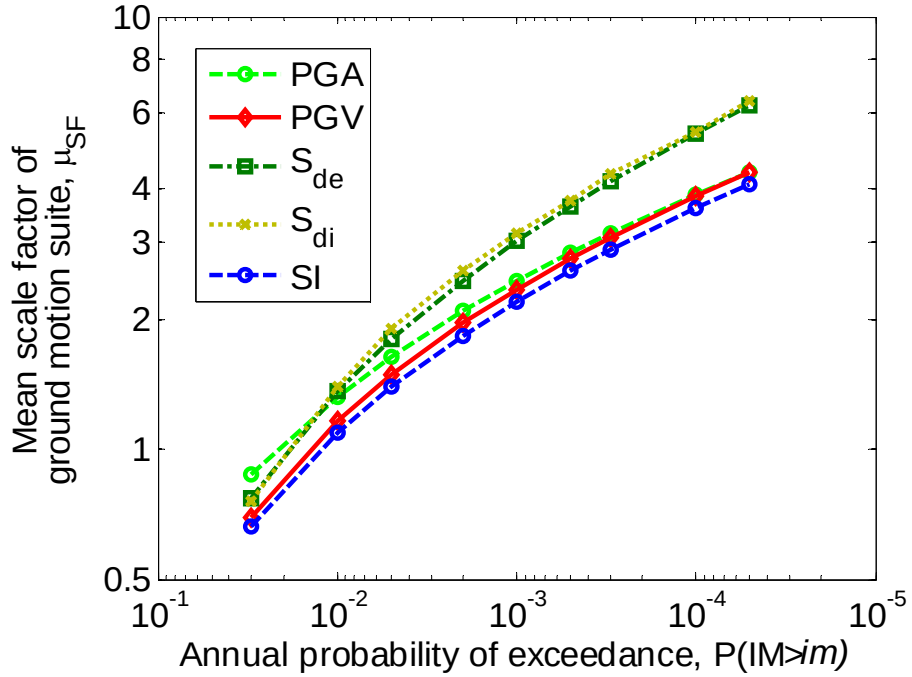


Figure 6-4: Mean scale factors of ground motion suite using various IMs.

Table 6-2: Details of the collapse capacity of the structure using different IMs.

IM	Median, $\overline{IM \mid C}$	$\sigma_{\ln IM \mid C}$	$P(C \mid IM_{2\mid 50})$*	P_C (x10^{-4})
PGA (g)	1.45	0.70	0.18	5.0
PGV (cm/s)	107	0.44	0.10	2.1
S_{de} (cm)	43.3	0.38	0.21	3.8
S_{di} (cm)	38.3	0.39	0.24	5.2
SI (cm.s/s)	425	0.41	0.06	1.15

*IM= $IM_{2\mid 50}$ has a 2% (Poissonian) probability of exceedance in 50 years.

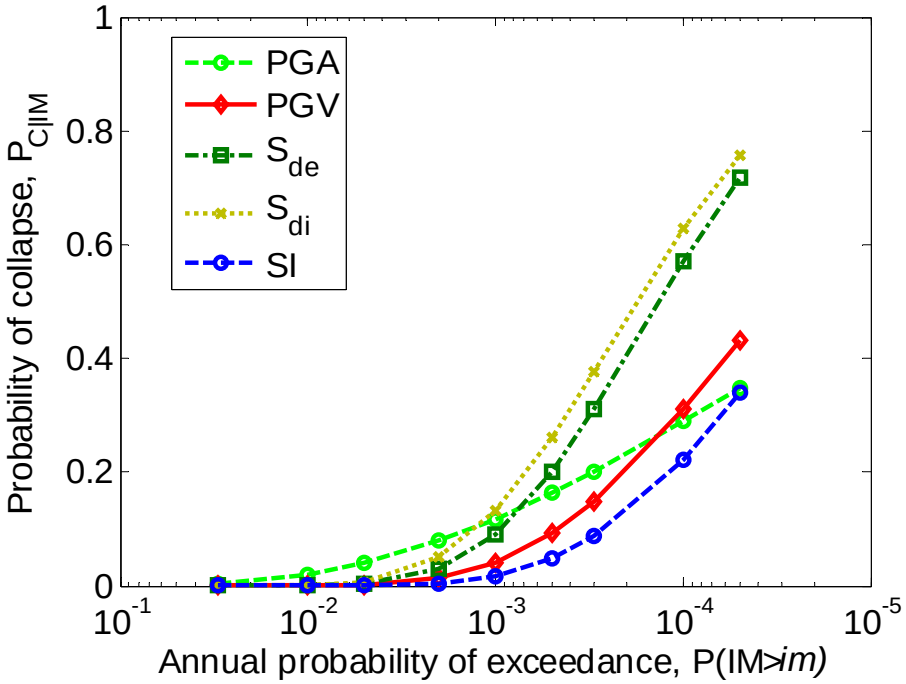


Figure 6-5: Probability of structural collapse as a function of IM equivalent annual exceedance probability.

6.5.2 Seismic demand hazard

The seismic demand hazard can be computed in a similar manner to the collapse hazard, by combining the distribution of structural response for a given ground motion intensity, $P(EDP > edp|IM = im)$, with the seismic hazard curve:

$$P_{edp} = \int_{IM} P(EDP > edp|IM = im)\left|\frac{P(IM > im)}{dIM}\right|dIM \tag{6-14}$$

where P_{edp} is the probability of exceeding EDP = *edp*.

Figure 6-6 illustrates the demand hazard curves for floor acceleration and interstorey drift ratios at both lower and upper floors in the case study structure. There is a negligible difference in the demand hazard curves using the various IMs at small EDP levels. It should be noted that the asymptotic values of the demand hazard are equal to the annual collapse probability, P_c, given in Table 6-2. For EDP values which have an exceedance probability of less than 10^{-2}, the difference in the demand hazard curves for the different IMs becomes evident. As the demand hazard depends on both the structural response distribution and ground motion hazard, then both IM efficiency (i.e. uncertainty in the EDP|IM distribution) and IM predictability (i.e. affecting the seismic hazard curve, $P(IM > im)$) are important.

Both of these two aspects are clearly evident in Figure 6-6a which illustrates that for the 2nd floor peak acceleration, the demand hazard computed using *PGA* lies below that of *PGV*, S_{de} and S_{di}, for accelerations up to approximately 1.0g because of its high efficiency [4, Figure 5]. Similar results for the maximum roof acceleration (Figure 6-6b) are observed as in the case of the Figure 6-6a, except that the demand hazard using *PGA* (relative to the other IMs) is increased because it is less efficient at predicting accelerations in upper floors, which contain significant modification from the characteristics of the structural vibration. Figure 6-6c illustrates the 3rd-4th floor interstorey drift ratio hazard. It is clearly seen that using *PGA* results in the largest hazard because of its poor efficiency, while *PGV* and *SI*, both give lower demand hazards than S_{de} and S_{di}, because of their superior predictability [4, Table 3] despite having slightly lower efficiency [4, Figure 4]. The maximum 10th floor-roof interstorey drift ratio hazard illustrated in Figure 6-6d shows similar trends as in Figure 6-6c.

All of the four plots in Figure 6-6 illustrate that the velocity-based IM's (i.e. *SI* and *PGV*) are the preferred IM's in terms of reducing the demand hazard, due to their high predictability despite not having the lowest efficiency. The theoretical reason for the relative importance of efficiency and predictability in reducing the demand hazard (as well as the loss estimation results to follow) is given in the discussion section later in the chapter.

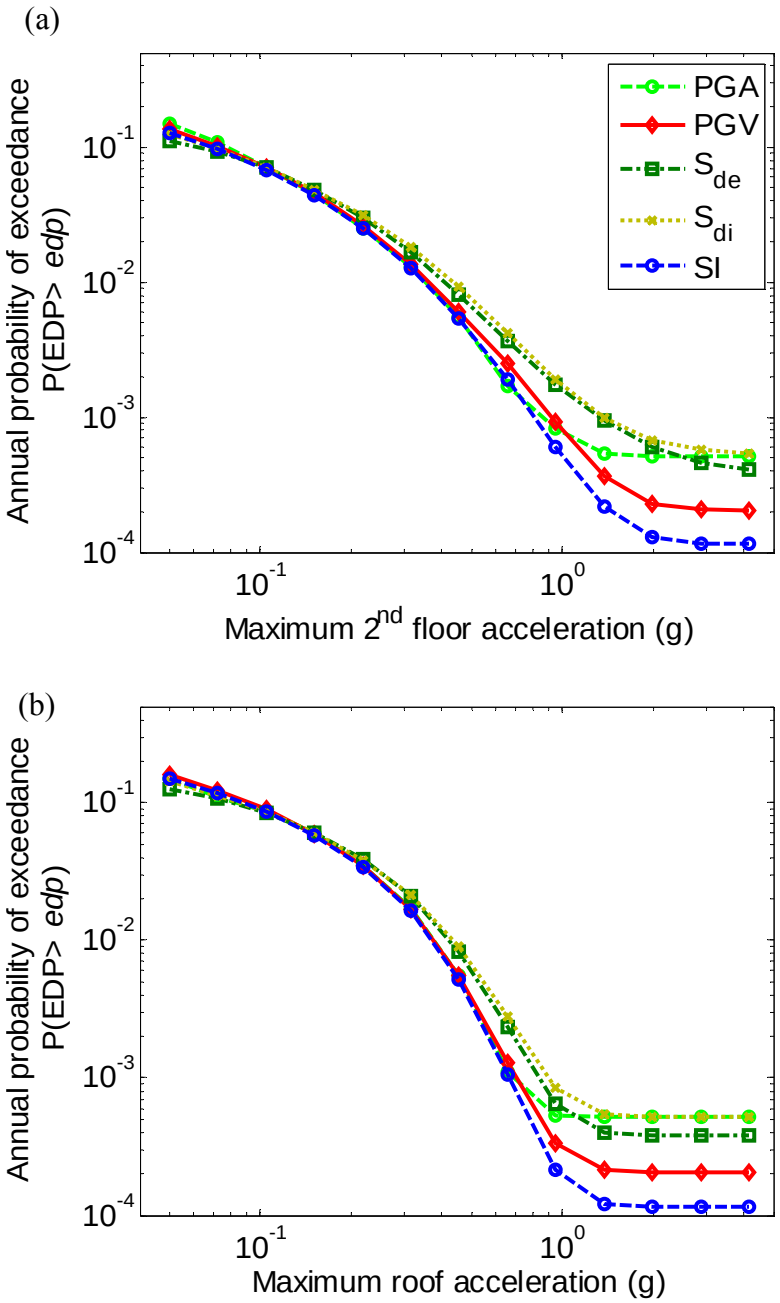


Figure 6-6: Demand hazard curves for: (a) maximum 2nd floor acceleration; (b) maximum roof acceleration; (c) maximum 3rd-4th floor interstorey drift ratio; and (d) maximum 10th floor-roof interstorey drift ratio.

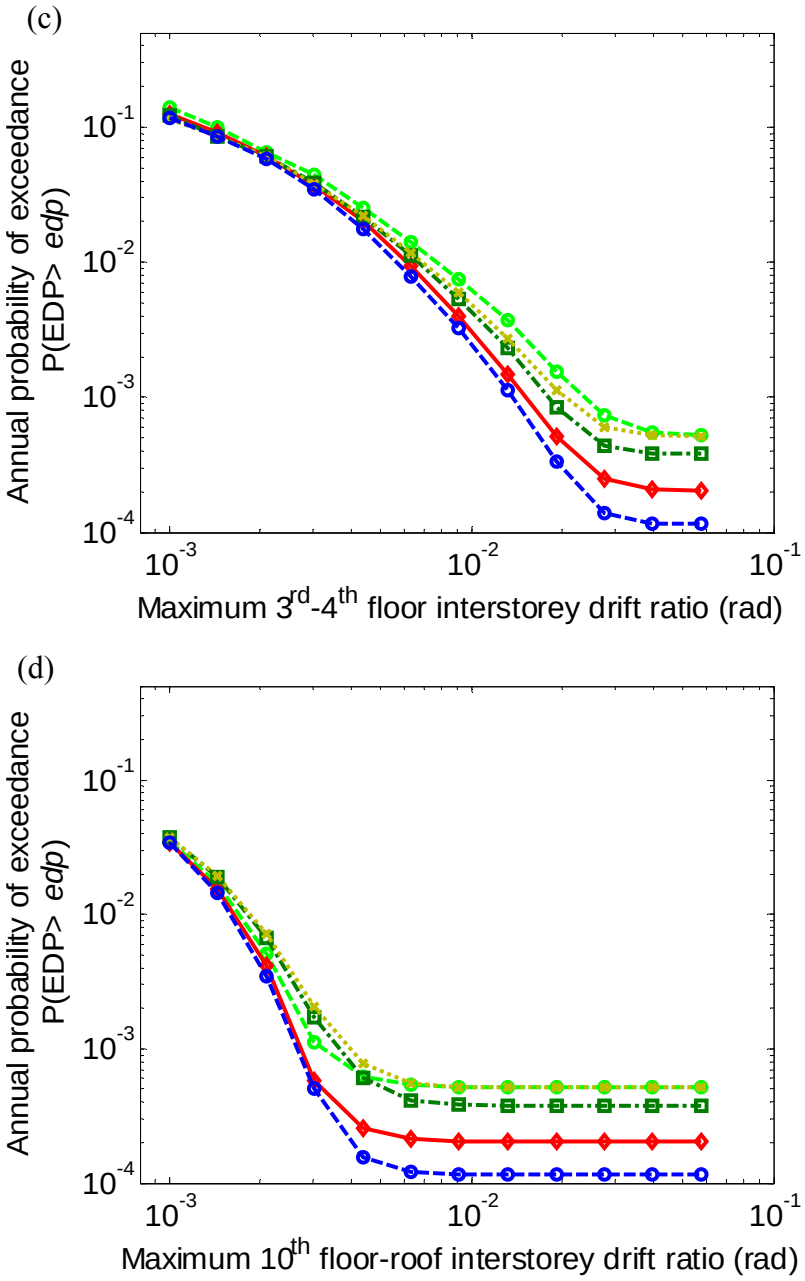


Figure 6-6 continued.

6.6 Seismic loss estimation

The ultimate goal of a seismic performance assessment is to estimate the consequences/losses to the entire structure, contents and occupants due to the occurrence of an earthquake event. Seismic loss estimation methods can be used to quantitatively assess seismic performance by considering the component inventory of the structure, and their vulnerability. Here the effect of IM selection on seismic loss estimation of the case study structure is illustrated. Only direct repair losses are considered, owing to their maturity

compared to the consideration of other losses such as human casualties and business disruption. The effects of economic loss amplification ('demand surge') and cumulative damage due to aftershocks are not considered.

When conducting a loss assessment of a structure, it is important to consider all of the components which have the potential to significantly contribute to the loss due to earthquakes causing a wide range of shaking intensities. Taghavi and Miranda, [13] illustrate that structural, non-structural and contents components are significant contributors to the total cost in office, hotel and hospital buildings. The components considered in the case study structure are listed in Table 6-3. It was assumed that all of the contents and non-structural components were equally distributed over the height of the building with the exception of the roof mounted equipment (located on the roof only), server and network equipment (located on the 3rd, 6th, and 10th floors), and elevators (ground floor). Mathematical details which are required to perform the loss assessment can be found elsewhere [1, 2].

6.6.1 Mean and standard deviation of loss given intensity

Figure 6-7a illustrates the expected loss given intensity (at equivalent probabilities of exceedance obtained from seismic hazard curves [4]) computed for the case study structure using the five IMs. It is noted that when shown in logarithmic scale the comparative trend between the different IM's is very similar to the mean scale factor applied to the ground motion suite as shown in Figure 6-4. The expected loss for a given IM depends on both the predictability of the IM (giving the equivalent probability of exceedance); the efficiency of the IM at predicting the spatially distributed demands in the structure; and the uncertainty in the damage states of the components in the structure [1, 2] (which are obviously independent of the choice of IM). In Figure 6-7 it is clear that the difference between the IMs in terms of predictability is more significant than the difference in efficiency, with the hierarchy of the IMs closely related to the predictability (e.g. as shown by the similarity of Figure 6-4 and Figure 6-7a). It should be noted that the reason for the reduction in the expected loss with increasing intensity based on PGA, relative to other IMs, is because of the large uncertainty in the collapse fragility curve using PGA. This large uncertainty means that it is more likely to have collapse at small IM levels (increasing the total loss), but less likely (compared to other IMs) at smaller exceedance probabilities (e.g. Figure 6-5).

In addition to the expected loss, uncertainty in loss given intensity is also important. The uncertainty depends on the correlations between demand, damage, and loss for different

components. Figure 6-7b illustrates the lognormal standard deviation (dispersion) in the loss given intensity (for an equivalent exceedance probability) when perfect correlations are assumed between demand, damage and loss in different components. As the ground motion intensity increases the dispersion in the loss reduces (while the expected loss increases), as also noted by Krawinkler [14, Figure 6.2]. Hence, for a given probability of exceedance (obtained from a seismic hazard curve) those IMs which give the lowest expected loss will consequently give a higher dispersion in loss (which is observed when comparing between the different IM's in Figure 6-7a and 6-7b).

Table 6-3: Components and quantities used in the case study loss estimation

Component	Description	Quantity	Reference
Ductile beam-column joints	Post 1960s ductile beam column joints (2 beams)	24 / floor	Williams *et al.* [15]
Columns	Gravity columns (and seismic columns on first floor)	20 on 1st floor, 4 on all other floors	Williams *et al.* [15]
Slab-beam-column connections	Connection of slab to seismic frame	24 / floor	Aslani [1]
Partition	Drywall partitions and finish	721 m^2 / floor	Aslani [1], ATC-58 [16], Porter *et al.* [17]
Exterior glazing	1.5m x 1.8m standard glass panes	99 panes / floor	ATC-58 [16]
Acoustical ceiling	0.6m x 1.2m tiles with Aluminium frames	693 tiles / floor	ATC-58 [16], Porter *et al.* [17]
Automatic sprinklers	3.7m sections of sprinkler piping	23 sections / floor	Porter *et al.* [17]
Servers and network equipment	Typical	\$260,000 on floors 3,6, and 10	ATC-58 [16]
Computers and printers	Typical	\$93000 / floor	ATC-58 [16]
Bookcases and file cabinets	Typical	\$16200 / floor	ATC-58 [16]
Roof mounted equipment	Coolers, airconditioning etc.	\$600,000 on roof	ATC-58 [16]
Workstation desks	Typical	\$21600 / floor	Buchan [18]
Generic acceleration sensitive	fire protection systems, HVAC, Heating, cooling, pumps, plumbing, toilets	\$100,000 / floor	Aslani [1]
Generic drift sensitive	vertical piping, bath tubs, F.H.C, Ducts	\$100,000 / floor	Aslani [1]

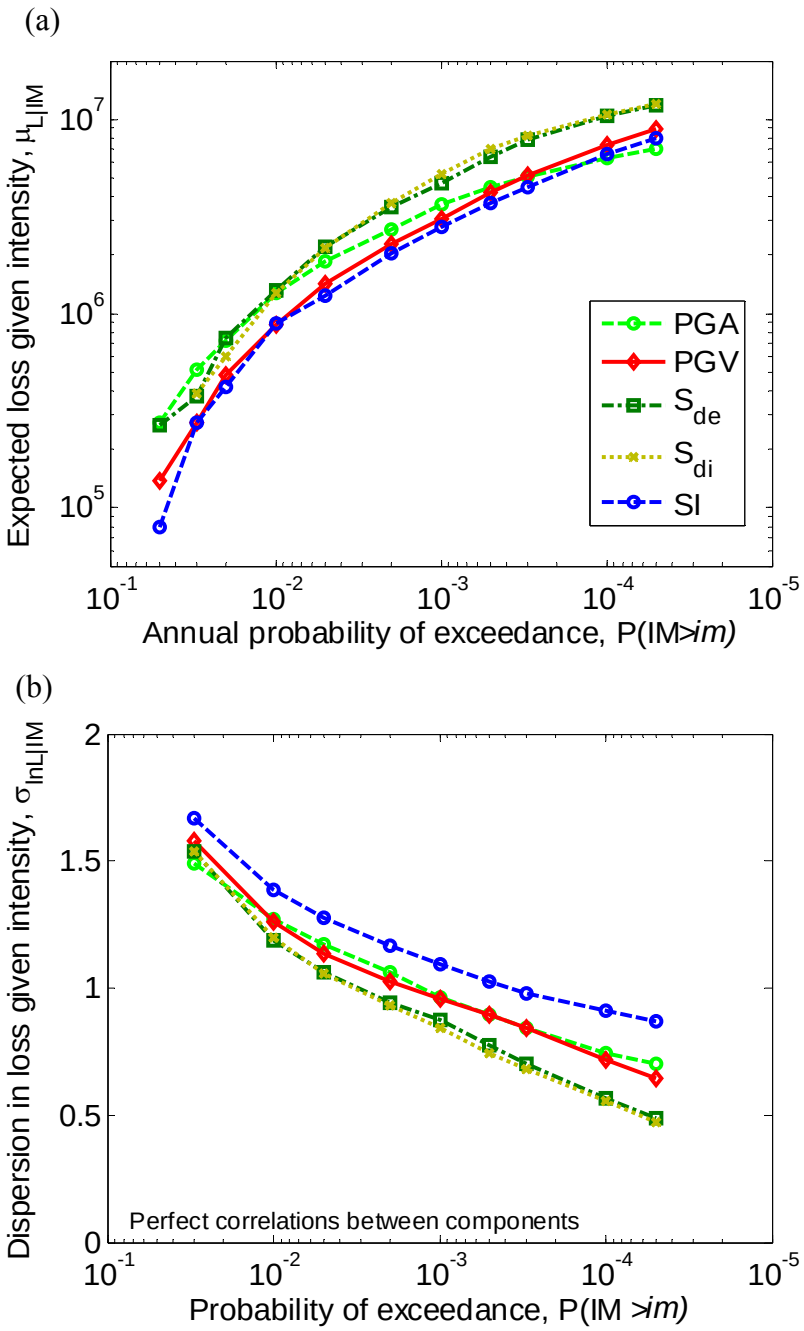
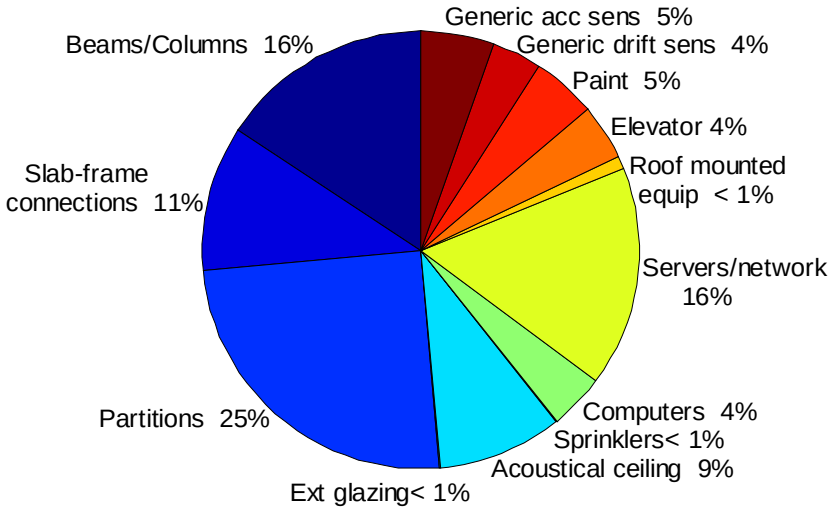


Figure 6-7: Loss given intensity measure at equivalent exceedance probabilities: (a) expected loss; and (b) dispersion in loss (perfect correlations).

While predictability of an IM appears to be the dominant effect from the preceding paragraphs, the efficiency of the various IMs, however, affects the distribution of the loss for a given intensity. This effect is particularly noticeable in terms of the contribution of acceleration- and displacement-sensitive components. Figure 6-8 illustrates the deaggregation of the expected loss given intensity (which has a 1/475 probability of exceedance) using *PGA* and S_{de}. Figure 6-7 illustrates that at this probability of exceedance, the expected loss using the two IMs is similar. As *PGA* is efficient at predicting acceleration demands, but inefficient

at predicting displacement demands (and vice versa for S_{de}), then the deaggregation of the expected loss using PGA indicates a reduction in the proportion of losses due to acceleration-sensitive components (e.g. acoustical ceiling, computers, servers/network, roof mounted equipment, elevator) compared to S_{de}. Similarly, using S_{de} results in a reduction of the proportion of losses due to drift-sensitive components (e.g. beams/columns, slab-frame connections, partitions, paint) compared to *PGA*.

(a)



(b)

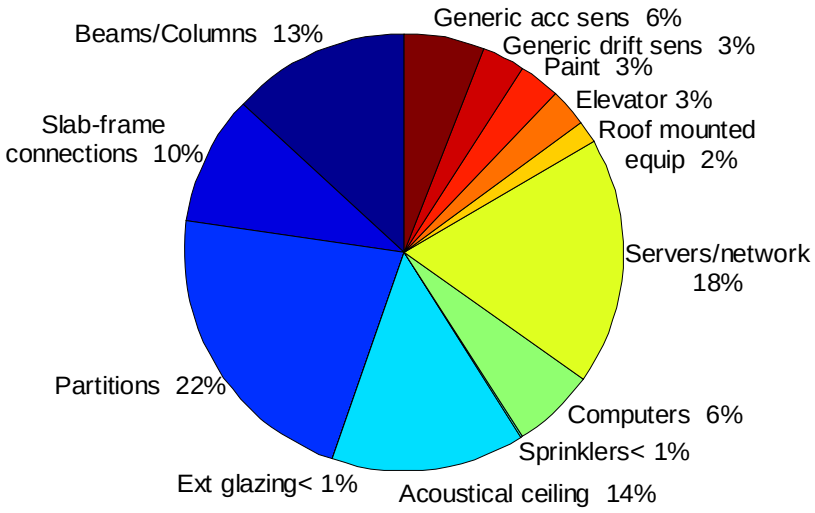


Figure 6-8: Deaggregation of expected loss at $P_{IM>im}$ = 1/475 using: (a) IM = *PGA*; and (b) IM = S_{de}.

6.6.2 Loss hazard

The loss hazard is computed by integrating the distribution of loss for a given intensity (defined by the mean and standard deviation discussed above) and the ground motion hazard curve and gives the probability of exceeding a specified level of direct repair loss (i.e. excluding casualties and business disruption). Figure 6-9a and 6-9b illustrate the resulting loss hazard curves for the cases of zero and perfect component correlations, respectively. With either assumption it can be seen that using *SI* and *PGV* give the smallest probability of exceedance for a given loss or vice versa. This result is due to the high predictability of *SI* and *PGV*, and the fact that they are also relatively efficient at predicting the spatially distributed acceleration and displacement demands in the case study structure [4]. The relative hierarchy of the other IMs is dependent on the value of loss being considered, with PGA being the highest over small loss values, but then dropping below S_{de} and S_{di} at higher loss values. The reason for the relatively poor performance of S_{de} and S_{di} is that they are have a poor predictability (relative to the other IMs), and despite providing high efficiency for peak interstorey drift ratios in the central portion of the structure (where the maximum drift ratios occur over the height of the structure), they are inefficient at predicting peak floor accelerations, and interstorey drifts in upper floors which are dominated by higher vibration modes.

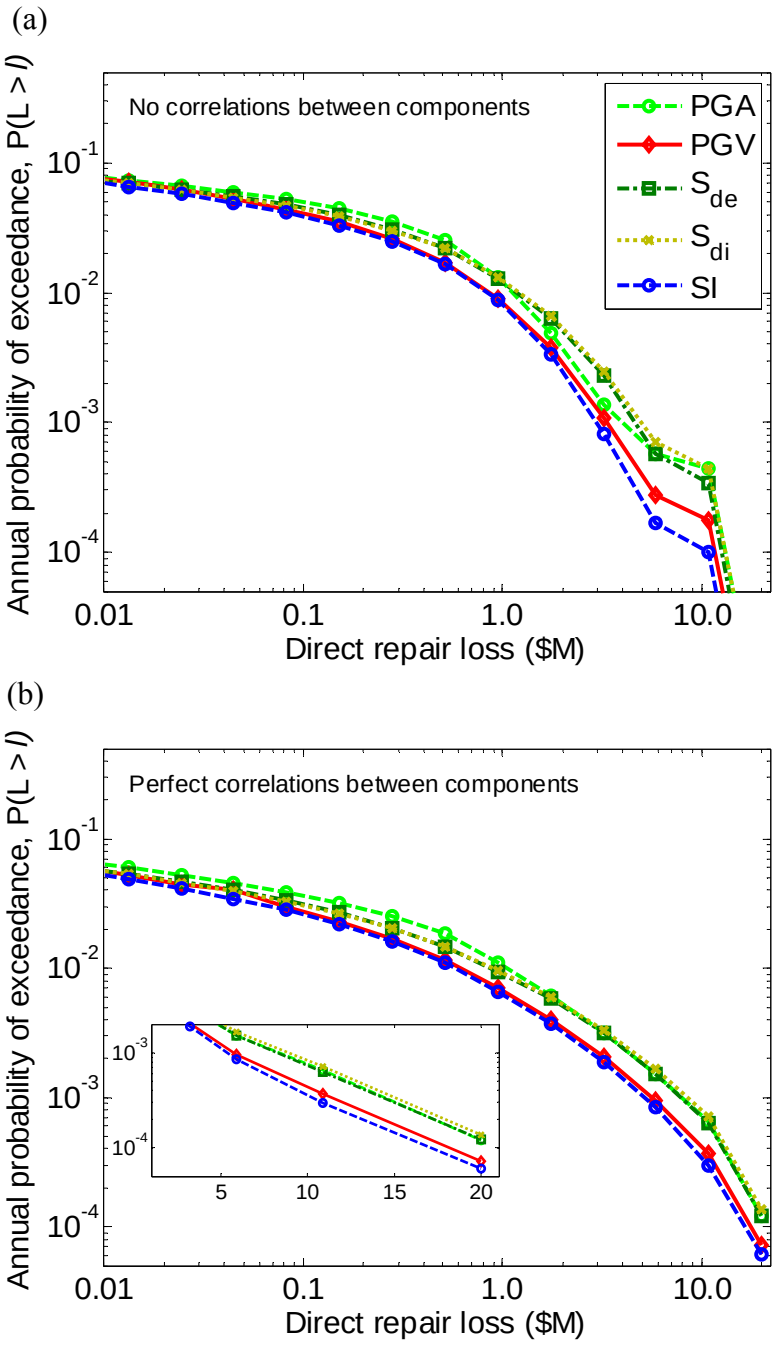


Figure 6-9: Loss hazard curves for: (a) no correlations between components; and (b) perfect correlations between components.

6.7 Discussion

Appendix B illustrates how both uncertainty in the prediction of ground motion (predictability) and seismic response (efficiency) affect the computation of the demand hazard. The fact that the total uncertainty is a square-root-sum-of-squares form means that the larger of the two uncertainties will dominate. Consider the prediction of the 3rd-4th floor interstorey drift ratio of the case study structure. Using S_{de} the predictability is 0.69 [4, Table

3]; while the efficiency is approximately 0.36 [4, Figure 4]; giving a total uncertainty of $\sqrt{0.69^2 + 0.36^2} = 0.78$. Alternatively *SI* has a predictability of 0.59 [4, Table 3] and an efficiency of approximately 0.43 [4, Figure 4] giving a total uncertainty of 0.73. Because the uncertainty in the ground motion prediction is typically larger than that of the seismic response uncertainty, then it is a more dominant uncertainty in the computation of the demand hazard (as well as loss estimation computations)

The consideration of the importance of both efficiency and predictability can be used to explain the majority of the observations presented in this chapter (and others for that matter). For example, while inelastic spectral displacement would be expected to give a better prediction than elastic spectral displacements for structures in the non-linear range (i.e. more efficient), Figure 10 of Tothong and Cornell [19] illustrates that as the extent of non-linearity increases (quantified by an R factor) the uncertainty in the prediction of S_{di} increases up to dispersions of 0.76. These two differences (i.e. improved efficiency, but reduced predictability) basically negate each other to give the similar demand hazards presented herein. The same argument as above can be made between the use of *PGA* and S_{de}. Because of its sufficiency and efficiency in predicting peak interstorey drift ratios many have made the comment that S_{de} is better that *PGA* (e.g. [20]). It is the demand hazard, however, which should be used for comparing the repercussions of IM selection, as it provides a result which depends on IM efficiency, sufficiency and predictability.

Somewhat of a disclaiming statement is warranted in regard to the previous paragraph. The authors are not suggesting one should simply select those intensity measures which are the most predictable. Recall that bias in the predicted distribution of EDP (sufficiency) was strongly correlated to efficiency of the IM for predicting the EDP. If one is trying to estimate the response of an elastic single-degree-of-freedom system it is unquestionably better to use S_{de} (which is perfectly efficient and sufficient) than using PGA and having to take care in selecting ground motion records to obtain (practically) the same result. What the authors are simply suggesting is that there is a negative impact on performance assessment if an IM is selected which is difficult to predict and does not strongly relate to the seismic demand being predicted. An example from this chapter is the use of S_{de} to predict peak accelerations. In this case S_{de} is both less efficient and less predictable than *PGA* resulting in a higher demand hazard. As *PGA* and S_{de} give similar hazards for displacement demands then it is the over-prediction of the acceleration demands that causes S_{de} (and S_{di} for that matter) to result in higher loss estimation results (Figures 6-7 and 6-9).

An obvious improvement on the work presented here would be the use of a vector-valued intensity measure [10], comprising scalar intensity measures which effectively predict high and low frequency dominated seismic responses (accelerations and displacements). Further to this, there is now an emergence of structure-specific demand prediction relationships, where the structure in consideration is subjected to hundreds if not thousands of ground motion records covering a range of magnitude and distance ranges (e.g. [21]). Regression using mixed-effects models can then be performed on the seismic demand (as opposed to simply the seismic intensity as is done to develop ground motion prediction equations). The problems with this approach are obviously the large (but ever-reducing) computational demand; correct selection of the database of records to use (making sure 'host' records are capable of being observed at the 'target' site of the structure); expertise to carry out the regression; and handling of regression equations in the M_w, R space poorly constrained by data. Furthermore, if this approach is to be extended to loss estimation, then uncertainties in the component fragility and loss will require regression of uncertain losses for each ground motion in the database, further complicating the regression procedure.

6.8 Conclusions

This chapter has examined the effect of intensity measure (IM) selection on the results of a seismic performance assessment of a 10 storey RC frame structure. The intensity measures examined were: peak ground acceleration, *PGA*; peak ground velocity, *PGV*; elastic and inelastic spectral displacement, S_{de} and S_{di}; and spectrum intensity, *SI*.

A method, based on multivariate regression, was proposed to account for the dependence of seismic response on parameters such as moment magnitude, M_w, source-to-site distance, R, and epsilon, ε, such that the same ground motion suite can be used over a range of different ground motion intensity levels.

Comparison of demand and loss hazard curves using the various IMs illustrates that the uncertainty in the ground motion prediction equation (predictability) typically is more significant than uncertainty in the seismic response prediction (efficiency). Most notably, results illustrate that the conventional IM, spectral displacement at the first mode, $S_{de}(T_1)$, can predict peak drift demands well, but its poor prediction of peak accelerations leads to higher loss estimates than alternative velocity-based IM's, namely spectrum intensity, *SI*, and peak ground velocity, *PGV*.

The structure used to obtain the loss estimation results presented had a component

inventory such that the losses due to drift- and acceleration-sensitive components were both significant in contributing to the total loss. Clearly, if the structure had a different purpose, i.e. a hospital, then the total loss would primarily be comprised by damage to medical equipment which will most likely be acceleration-sensitive. In such a case, *PGA* is likely to be the optimal intensity measure for use in loss estimation.

6.9 Acknowledgements

Financial support of the first author from the New Zealand Tertiary Education Commission Bright Futures scheme is appreciated.

6.10 References

[1] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[2] Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, and Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society for Earthquake Engineering* 2008.

[3] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1973-1997, DOI: 10.1002/eqe.694.

[4] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: ground motion and structural response. *Earthquake Engineering and Structural Dynamics* 2008; (submitted).

[5] Bazzurro P. Probabilistic seismic demand analysis. Stanford University, Stanford, CA, 1998. 329. http://www.stanford.edu/group/rms/

[6] Haselton CB. Assessing Collapse Safety of Modern Reinforced Concrete Moment Frame Buidlings. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2007, 312.

[7] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[8] Kutner M, Nachtsheim C, Neter J, and Li W. *Applied Linear Statistical Models*. Fifth Edition ed. McGraw-Hill/Irwin: New York, 2005; 1396.

[9] Hastie T, Tibshirani R, and Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer: New York, 2001; 533.

[10] Baker JW. Probabilistic structural response assessment using vector-valued intensity measures. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1861-1883.

[11] Baker JW and Cornell CA. Spectral Shape, record selection and epsilon. *Earthquake Engineering and Structural Dynamics* 2006; **35**(9): 1077-1095.

[12] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1901-1914.

[13] Taghavi S and Miranda E. Response assessment of non-structural building elements. 2005. http://peer.berkeley.edu/publications/peer_reports.html

[14] Krawinkler H. Van Nuys hotel building testbed report: exercising seismic performance assessment. PEER 2005/11 Pacific Earthquake Engineering Research Centre, 2005. 264.

[15] Williams MS, Villemure I, and Sexsmith RG. Evaluation of seismic damage indices for concrete elements loaded in combined shear and flexure. *ACI Structural Journal* 1997; **94**(3): 315-322.

[16] ATC-58. Guidelines for seismic performance assessment of buildings: ATC-58 35% Draft. Applied Technology Council, 2007. http://www.atcouncil.org/atc-58.shtml (last accessed 19/11/08)

[17] Porter KA, Kiremidjian AS, and LeGrue JS. Assembly-Based Vulnerability for Buildings and Its Use in Performance Evaluation. *Earthquake Spectra* 2001; **17**(2): 291-312.

[18] Buchan J. Internal report of non-structural and contents inventories for the Christchurch Council Chambers building. Christchurch, New Zealand., 2007.

[19] Tothong P and Cornell CA. An Emperical Ground Motion Attenuation Equation for Inelastic Spectral Displacement. *Bulletin of the Seismological Society of America* 2006; **96**(6): 2146-2164.

[20] Shome N, Cornell CA, Bazzurro P, and Carballo JE. Earthquakes, records, and nonlinear responses. *Earthquake Spectra* 1998; **14**(3): 469-500.

[21] Hancock J, Bommer JJ, and Stafford PJ. Numbers of scaled and matched accelerograms required for inelastic dynamic analyses. *Earthquake Engineering and Structural Dynamics* 2008; **37**(14): 1585-1702, DOI: 10.1002/eqe.827.

7.Efficient Evaluation of Performance-Based Earthquake Engineering Equations

Bradley BA, Lee DS, Broughton R, Price C. Efficient evaluation of performance-based earthquake engineering equations. *Structural Safety* 2009; **31**(1): 65-74.

7.1 Abstract

In this chapter attention is given to the efficient numerical evaluation of the Pacific Earthquake Engineering Research (PEER) performance-based earthquake engineering framework equations. In particular, potential problems in determining an adequate yet efficient region of integration are discussed. An algorithm called "Magnitude-oriented Adaptive Quadrature" (MAQ) is developed, which is an integration algorithm with both locally and globally adaptive capabilities. MAQ allows efficient integration over the entire integration domain and requires only an error tolerance and maximum number of function evaluations to be specified. The advantages of utilizing the MAQ algorithm over other conventional integration methods such as Romberg integration and conventional adaptive quadrature are illustrated for the numerical computation of (1) expected annual loss; and (2) annual rate of collapse. It is shown that for determination of the expected annual loss a 4.5- to 8.8- fold reduction in the computational demand is obtained using MAQ compared to conventional integration methods. For annual rate of collapse the computational demand reductions range from 30% to two-fold. The computational reductions are a function of the error tolerance prescribed, with greater computational reductions as stricter tolerances are enforced.

7.2 Introduction

With the rapid growth of performance-based earthquake engineering in research and design, the accurate and efficient evaluation of the governing probabilistic integrals increases in importance. If the functions contained in the integrand are of an appropriate form then it may be possible to obtain a 'closed-form' analytical solution. Such analytical solutions have been presented in the literature, for example: annual rate of exceedance of demand [1-3] and annual rate of structural collapse [4, 5]. In general, these 'closed-form' solutions use functions which are local approximations to the true functions defining the arguments of the integral equations. When no simplification is possible without significant loss of accuracy, or the dimensionality of the problem becomes large, these integrals must be solved using a numerical scheme.

Numerical solution of integrals can utilize simulation (e.g. Monte Carlo) or a direct integration technique (e.g. quadrature). Some examples of simulation methods in PBEE include Assembly-Based Vulnerability [6] , and Subset Simulation [7]. Direct numerical integration is typically employed to evaluate relationships within the so-called PEER framing equation [8], advocated by the Pacific Earthquake Engineering Research (PEER) Centre. This is because the PEER framing equation makes the Markovian assumption allowing the triple integral to be computed as successive single integrals. It is well realised that unlike quadrature techniques, such simulation methods do not suffer from the 'curse of dimensionality' [9] making their use particularly suited to high-dimensionality problems. As attention herein will be given to solution of equations within the PEER framework formula, we will restrict our attention to direct numerical integration methods.

Baker and Cornell [10] propose the use of the first-order second-moment (FOSM) method as a potential method of uncertainty propagation for certain relationships with the PBEE framework. They note however that certain relationships (i.e. those including the ground motion hazard curve) should be used with direct numerical integration (as the ground motion hazard is likely a dominant contributor toward the total uncertainty in the seismic performance measure of interest). Also the FOSM method, being a first-order approximate method, has limitations which may make it significantly inaccurate in certain situations [11].

The use of performance-based frameworks within which these integrals are solved in a practical situation requires that efficient integration algorithms are used offering: (i) numerical efficiency; (ii) accuracy tolerance specifications; and (iii) 'user-friendliness'. Efficiency is of

primary importance since the computational demand will influence whether such an approach will be employed. Accuracy tolerance is important as a user should be able to specify a relative error tolerance, and know that the results of the analysis will be accurate to within that tolerance (with the exception of pathological cases). Here we have used the general phrase 'user-friendliness' to indicate the information that is required by the integration algorithm (such as integration region, step size and error tolerances), which is important in the avoidance of pitfalls for inexperienced users, and will be elaborated further on in the chapter.

Although many standard references are available on numerical solution algorithms, as will be shown in this chapter, it is possible to take advantage of the form of the integrand to develop an algorithm with significantly greater efficiency than 'general-purpose' algorithms, and which requires only the integration error tolerance and maximum number of function evaluations (and not any information on step size or integration region).

7.3 The risk equations

For brevity, throughout the remainder of this chapter the probabilistic integral equations which comprise the PEER framing formula will be referred to as 'the risk equations'.

There are several different appearances of the risk equations, the most well-known being the so-called 'Triple-integral formula' [8] (Equation (7-1)), which gives the annual frequency of exceeding some decision variable, *DV*.

$$\lambda(dv) = \iint \sum G(dv|ds)\left|\Delta G(ds|edp)\right|\left|dG(edp|im)\right|\left|d\lambda(im)\right| \tag{7-1}$$

where *im* = intensity measure (e.g. *PGA*); *edp* = engineering demand parameter (e.g. peak interstorey drift); *ds* = damage state; *dv* = decision variable; $G(x|y) = G(X \geq x|Y=y)$ is the complementary cumulative distribution function (CCDF) of x given y; $\lambda(z)$ is the annual frequency of exceeding z; $dG(x|y)$ and $d\lambda(z)$ are the differentials of $G(x|y)$ and $\lambda(z)$, respectively.

The triple integral formula can be de-coupled into successive single integrals based on the conditional independence assumption (e.g. in Equation (7-1), $G(DV|DM)$ is dependent only on *DS* and not on *EDP*, or *IM*). Such de-coupling can lead to for example, the rate of exceedance of some level of demand [1-3], and the expected loss given an intensity measure [12], which are given in Equation (7-2) and (7-3), respectively.

$$\lambda(edp) = \int G(edp|im)\left|d\lambda(im)\right| \tag{7-2}$$

$$E\left[L_T \middle| im\right] = \int E\left[L_T \middle| edp\right] \left| dG\left(edp \middle| im\right)\right| \tag{7-3}$$

where $E[X|y]$ = the expected value of X given $Y = y$; and L_T = total loss.

Since the conditional independence assumption allows a reduction to successive single integrals, then herein we will be concerned with integrals of the following general form:

$$I = \int_{all\ x} G(x)\left|dF(x)\right| = \int_{all\ x} G(x)\left|\frac{dF(x)}{dx}\right| dx = \int_{all\ x} G(x) f(x) dx \tag{7-4}$$

where $G(x)$ and $F(x)$ are piecewise-continuous functions; $dF(x)$ is the differential of $F(x)$; and $\frac{dF(x)}{dx} = f(x)$ is the derivative of $F(x)$ with respect to x. All of the functions comprising the integrand are functions of the integration variable, x. As far as the authors are aware, Equation (7-4) encompasses all of the PBEE equations presented in the literature, with the one known exception that of vector-based equations (e.g. [13, 14]), in which case $G(\boldsymbol{x})$ and $F(\boldsymbol{x})$ are vector-valued functions (i.e. $\boldsymbol{x} = [x_1, x_2, \ldots x_N]$) and the multiple integration is over the domain of $\boldsymbol{x}$. Such vector-based integrals will not be dealt with explicitly here, although the concepts presented herein are still applicable and are discussed briefly later in the chapter (as quadrature-based multi-dimensional numerical integration is effectively embedded single numerical integrations). As previously mentioned it should be noted that as the dimensionality of the integrals increases, the efficiency of simulation based procedures will increase relative to that of direct numerical integration (and will eventually become more efficient [9]). As far as the authors are aware, only 2-dimensional vectors have been proposed for the *EDP*|***IM*** relationship (e.g. [13, 14]), and therefore also 2-dimensional vectors for the $\lambda(\boldsymbol{IM})$ relationship, but all damage (i.e. *DS*|*EDP*) and loss (i.e. *DV*|*DS*) relationships are scalar (e.g. [12, 15]).

7.4 Region of integration

As the integration variable (i.e. *IM* or *EDP*) is defined over all the positive real numbers (i.e. $x \in [0, \infty)$), then the integration (strictly speaking) is over this entire domain of the integration variable. However, in practice the magnitude of the integral tends (usually rapidly) to zero at the endpoints of the integration region (a requirement for convergence of the integral as x tends to infinity, and a practical constraint at x equals zero), thus allowing a sub-region of this domain (instead of the entire domain) to be used to evaluate the integral. For example, in Equation (7-3), it is typically assumed that the demand conditioned on

intensity, (*EDP*|*IM*), relationship has a lognormal distribution [16, 17]. Thus the region of integration can be specified as a certain number of standard deviations either side of the mean demand for the given intensity.

Inadequate selection of the sub-region for integration will lead to underestimation of the value of the integral. It is therefore desired that determination of the integration region be obtained internally within a computational algorithm, thus preventing possible errors by inexperienced users (this is one objective in the previously defined 'user-friendliness'). In certain cases it is not a trivial task to determine the region of integration because of several reasons which are discussed in detail below.

Problems with selection of the integration region can occur when the distribution that is used to 'guess' an appropriate integration region does not conform well to the shape of the integrand. To explain this more clearly we refer to the computation of the mean annual frequency of exceeding some level of demand given in Equation (7-2). For this problem the mean and standard deviation of the *EDP*|*IM,* relationship are typically provided via structural analysis instead of *IM*|*EDP,* which requires the use of iteratively scaling ground motion records until a target *EDP* is achieved [18]. However, for solution of the integral the mean and standard deviation of the intensity given demand (*IM*|*EDP*) relationship is required in order to determine the appropriate intensity values to integrate over. If it is assumed that locally EDP and IM can be related by $EDP \mid IM = \ln(a) + b\ln(IM) + \beta_{EDP|IM}\varepsilon$ (where all terms are deterministic except ε which has zero expectation and unit lognormal standard deviation), then the mean and dispersion of the intensity given demand relationship can be given by $\mu_{\ln IM|EDP} = \left[\mu_{\ln EDP|IM} - \ln(a)\right]/b$, and $\beta_{\ln IM|EDP} = (\beta_{\ln EDP|IM})/b$, respectively [19]. Figure 7-1 shows the probability density function (pdf) of the intensity given demand relationship, the integrand of the demand hazard for $EDP = 0.05$ (where in this case the EDP is assumed to be the peak drift of the structure), and the normalised pdf of the ground motion exceedance (i.e. not the ground motion hazard). Here, seismic intensity and demand are measured using the 1.5 second elastic spectral acceleration ($IM = S_a(T=1.5s,5\%)$), and drift ($EDP = \theta$), respectively; the *EDP*|*IM* parameters are (a, b, $\beta_{lnEDP|IM}$) = (0.01, 1.5, 0.4); and the ground motion hazard is for Christchurch, New Zealand [20]. It is immediately evident that the two curves in Figure 7-1 are offset. If the density of the *IM*|*EDP* relationship is used to approximate the bounds for the region of integration then for a desired level of accuracy it will likely lead to lower and upper integration limits which are too large. Having a lower limit which is too large will potentially mean that a sub-region of the integration domain

which contributes significantly toward the value of the integral will be neglected, resulting in potential significant error. This problem can be resolved by integrating over a large number of standard deviations (i.e. about ±4 is sufficient for a integration tolerance of $1x10^{-3}$). However in this case, the upper limit will be quite large (using 4 standard deviations gives an upper limit of 8.49) and since the magnitude of the integrand becomes insignificant around an intensity of 5, then integration over this insignificant region (i.e. from intensities of 5-8.5) unnecessarily increases the computational demand.

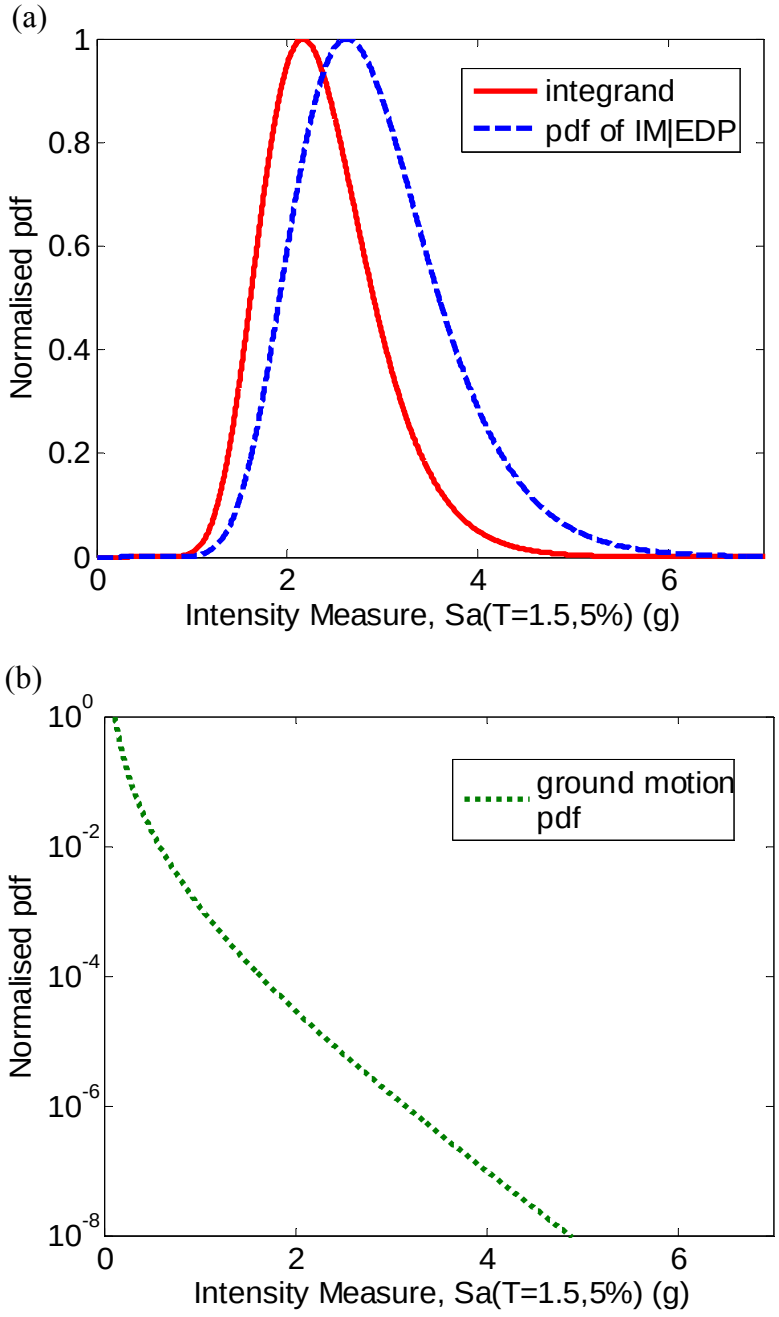


Figure 7-1: Illustration of the offset in the probability density function used for integration region estimation with: (a) the integrand of the integral; and (b) the ground motion pdf.

Problems in determining the region of integration can also occur when neither of the two functions comprising the integrand of the risk integral conform to an analytical probability density function (with the exception of extreme value-type distributions). This problem is encountered when trying to compute the expectation of the annual loss (Equation (7-3)). In this case the functions in the integrand are the expected loss given intensity and the derivative of the ground motion hazard, neither of which are a 'conventional' probability density. Figure 7-2 illustrates the expected loss given intensity for a typical NZ bridge [21]; ground motion hazard curve for Wellington, New Zealand [20]; and the resulting integrand of the expected annual loss. It can be seen that while it may be relatively simple to determine a sub-region of integration from inspection, trying to determine a sub-region within a computational algorithm will require the use of some optimisation algorithm (or similar) to determine where the values of the integrand are or are not significant. This would mean that numerous evaluations of the integrand would be required, none of which would be used toward computing the integral.

In the above two problematic cases it is desirable, instead of predefining a sub-region of integration, to instead integrate over the entire domain of integration (or for the first problem type, using a sub-domain based on a large number of standard deviations), but performing the integration in such a way that little computational effort is expended in computing the integrand over regions which contribute insignificantly toward the numerical value of the integral. An algorithm offering a potential solution to the above statements is addressed in subsequent sections.

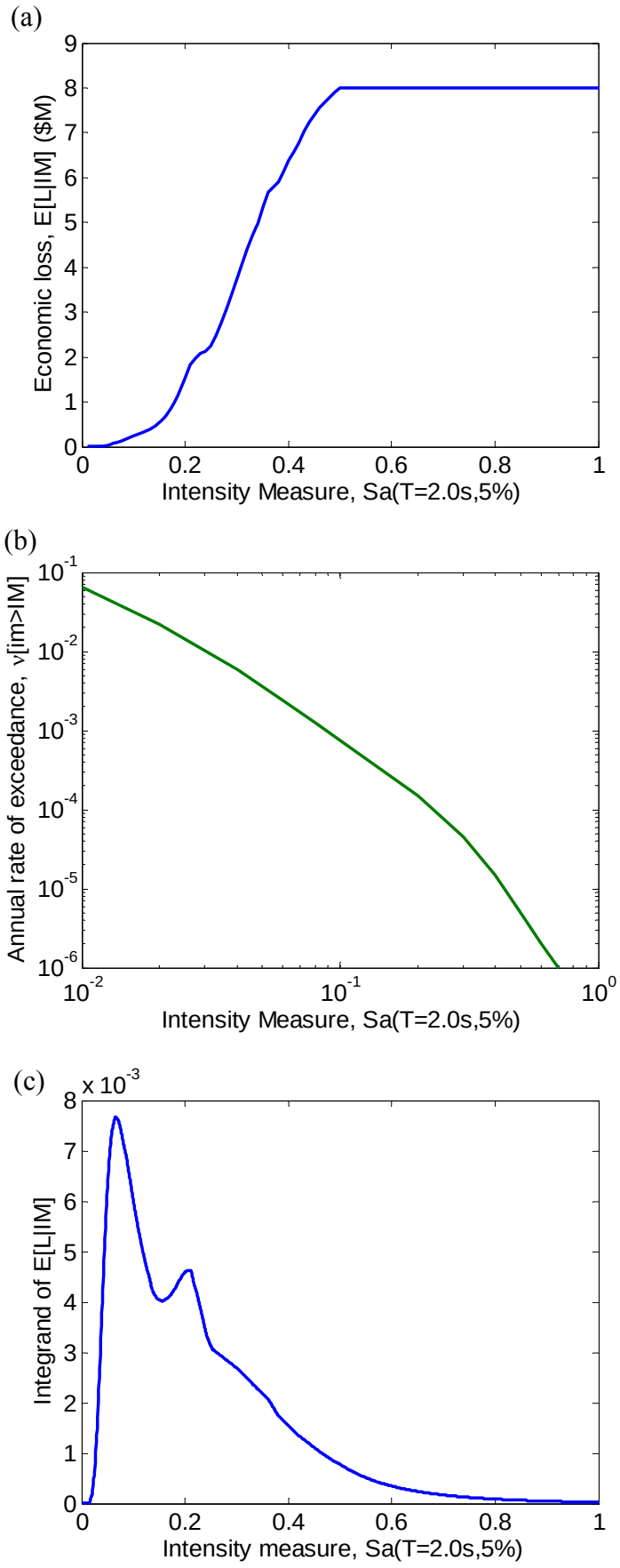


Figure 7-2: (a) Expected loss given intensity; (b) ground motion hazard; and (c) expected annual loss integrand for the example problem considered.

7.5 Indefinite limits of integration and assumptions

As the domain of x is not finite then Equation (7-4) cannot be easily numerically integrated in its current form. There are numerous methods in handling indefinite integrals, in this work the following mapping is used to make the integration domain finite:

$$t = \frac{1}{(x+1)} \tag{7-5}$$

This transformation maps the region x = [0, ∞) to t = (0,1], and in particular x = 0 → t = 1, x = ∞ → t = 0. With this transformation, Equation (7-4) can be re-written the following form:

$$I = \int_{x=0}^{x=\infty} G(x) f(x) dx = \int_{t=0}^{t=1} \frac{1}{t^2} G\left(1/t - 1\right) f\left(1/t - 1\right) dt \tag{7-6}$$

If the integrand, $G(x)f(x)$ approachs zero at least as fast as $1/x^2$ as x tends to infinity then the integral converges. This is the same as requiring $G\left(1/t - 1\right) f\left(1/t - 1\right) \to 0$ at least as fast as $t^2 \to 0$. Equation (7-6) is now in a form suitable for direct numerical integration.

We will also make two mild assumptions (which we will later relax) for the forms of the functions comprising the integrand. The first is that the function $G(x)$ is monotonically increasing, and the second that $f(x)$ is a uni-modal function (i.e. it has only one maximum, before which the function is monotonically increasing, and after which it is monotonically decreasing). These two assumptions generally agree with intuition, for example one would expect that as the level of ground shaking (intensity) increases, the level of damage (loss) increases, similarly the function, $f(x)$, is usually a probability density function which is assumed to be lognormal (e.g. the EDP|IM relation) or is similar to an extreme value distribution (e.g. $\lambda(IM)$), both of which are uni-modal functions. Under the above two assumptions it can be shown that while the integrand of I, $G(x)f(x)$, is not strictly a uni-modal function (although in the majority of cases it will be), the majority of the integral will be contributed from a single sub-region (e.g. Figure 7-1 and 7-2c), and the integrand will not have a significant multi-modal shape (i.e. a significant contribution toward the integral will not come from two or more distinctively separate regions of the integration variable). The premise that the dominant contribution toward the integral will occur over a single region of integration is used to target the computational effort toward the integral evaluation at the location of this region. Nevertheless, the proposed algorithm will cope with significant contributions from several distinct regions provided the initial distribution of integration points is capable of identifying each such region.

7.6 Concept of Magnitude-oriented Adaptive Quadrature (MAQ)

As previously stated, we aim to have a quadrature method which can adapt around the integrand in order to achieve both accuracy, computational efficiency, and require no integration computation specifics (such as the step size and region of integration), other than the error tolerance and maximum number of function evaluations. For completeness, we first discuss two simple quadrature methods, Romberg integration, and Simpson's rule-based adaptive quadrature, which are globally and locally adaptive methods, respectively. These two methods are described briefly here as they are used as a basis of the locally and globally adaptive quadrature method we propose. Further information on these two algorithms can be found, for example, in [22, 23].

One simple but highly efficient algorithm for numerical integration that allows error estimation is Romberg Integration [22]. The Romberg integration algorithm is a computational method of using Richardson Extrapolation with the Trapeziodal rule [23], which uses two approximations (of, in this case, an integral) to compute a third more accurate approximation. The Romberg Integration algorithm can be expressed in the following form:

$$I_{j,k} = I_{j+1,k-1} + \frac{1}{4^{k-1}-1}\left(I_{j+1,k-1} - I_{j,k-1}\right) \tag{7-7}$$

where k (>1) is the order of the approximation error (i.e. k = 2 corresponds to $O\left(h^4\right)$, k = 3 to $O\left(h^6\right)$ etc.; h is the step size of integration; the subscript j is used to denoted the more and less accurate approximations of the integral ($I_{j+1,k-1}$ and $I_{j,k-1}$, respectively); and $I_{j,k}$ is the improved estimate of the integral. Hence Equation (7-7) shows that effectively the improved estimate is obtained by incrementing the more accurate of the two estimates. Romberg integration is efficient in the sense that it is able to use previous integration points when the integration step size is further refined and then combined with Richardson Extrapolation it is markedly more efficient than conventional Simpson's rule [23].

The Romberg integration algorithm is most efficient when the curvature of the integrand is relatively constant over the region of integration. When the curvature of the integrand varies significantly over the region of integration, the rate of convergence is reduced because certain sub-regions take longer to achieve convergence (to a specified tolerance) while additional function evaluations are 'wasted' in sub-regions where convergence has already been achieved. The reduction in convergence of the Romberg algorithm in such cases is due to the fact that convergence of the integral is measured globally (i.e. over the entire region of

integration).

Adaptive quadratures [22] are one such set of algorithms where convergence is measured locally (i.e. over some sub-region), which makes it more efficient for integrals with rapidly changing curvatures. Adaptive quadrature was initially developed for use in solving ordinary differential equations where so-called 'stiff' problems [22, 23] are encountered. The most common form of adaptive quadrature is that which uses successively refined approximations of Simpson's rule. Simpson's rule-based Adaptive quadrature (herein referred to simply as adaptive quadrature for brevity) is based on the following steps: (1) numerically integrate over some sub-region using Simpson's rule with three points; (2) sub-divide the interval by evaluating the integrand at the ¼ and ¾ points of the sub-region, then apply Simpson's rule using three points over each of the two half-regions and then add the resulting two values; (3) compare the relative error between the more and less accurate approximations; if the relative error is acceptable then move to the next sub-region, else repeat the process, but using only the left half the original sub-region.

From the previous discussions the following points can be noted for Romberg integration and adaptive quadrature concerning convergence and efficiency. Romberg integration measures convergence globally and therefore efficiency can be reduced when certain regions of the integration domain take longer to converge than others. Adaptive quadrature measures convergence locally, therefore overcoming the aforementioned problem of local convergence for Romberg integration. However, because convergence is only measured locally then significant computational effort can be spend evaluating regions of the integration domain whose magnitude does not affect the global convergence. Below an algorithm, herein referred to as magnitude-oriented adaptive quadrature (MAQ), is introduced which aims to provide a solution to both of the noted problems in Romberg integration and adaptive quadrature, which in particular occur in the problem context discussed in this chapter.

MAQ can be thought of as a further extension of conventional adaptive quadrature. The key differences being that: (i) MAQ considers convergence at both a local and global level; and (ii) MAQ uses a modified region discretization. A schematic illustration of the four-step process of MAQ is given in Figure 7-3 and outlined in the following paragraphs. It is also noted that a MATLAB implementation of the MAQ algorithm can be obtained from http://www.civil.canterbury.ac.nz/postgrads/bbradley.asp. Also herein for brevity the region of integration is implied to be the region of integration after transformation by Equation (7-5) unless otherwise stated.

Initialisation: To start the algorithm requires the user to specify values for the region of integration (i.e. for x=(0, infinity) t=(0, 1)) for the error tolerance, 'tol', and maximum number of function evaluations, 'maxeval'. The initial value of the integral is set to zero as well as counter variables which are used to keep track of the number of function evaluations, and to store computed values in arrays.

Step 1: firstly the integrand is evaluated at the two endpoints (*a* and *b*) and a midpoint *c* = (*a*+*b*)/2. Based on these three points Simpson's rule is used to estimate the integral over the region of integration (I_1) which is given by:

$$I_1 = \frac{(b-a)}{6}\left[fa + 4\,fc + fb\right] \tag{7-8}$$

where *fa, fb* and *fc* denote the value of the integrand at *a*, *b* and *c*, respectively.

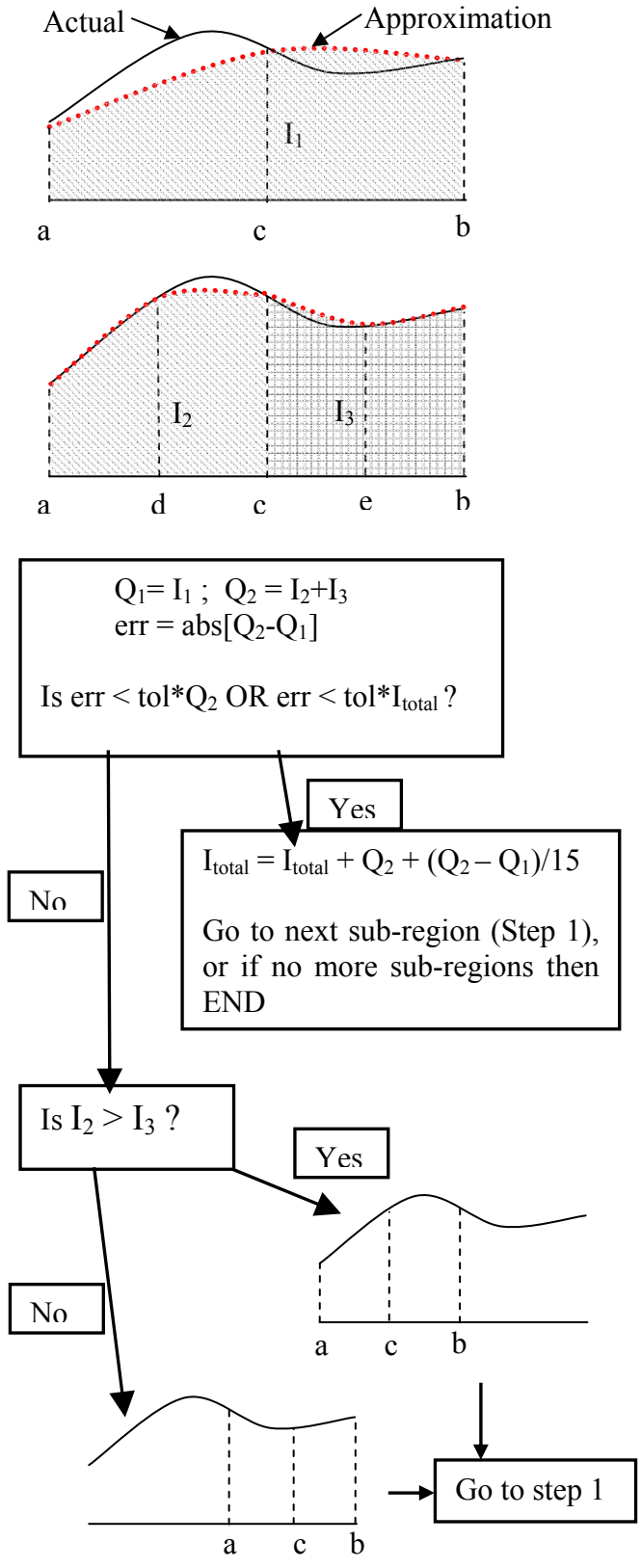
Step 2: Two additional integrand evaluations are obtained at the points, *d* = (*c*+*a*)/2 and *e* = (*b*+*c*)/2 (which are denoted as *fd* and *fe*, respectively). Thus, Simpson's rule (i.e. Equation (7-8) with the necessary subscript substitutions) can be used to obtain two integral approximations over the regions (a, c) and (c, b). As a smaller step size of integration is used these integral approximations (I_2 and I_3) will be more accurate than the approximation obtained in step 1 (I_1).

Step 3: An approximation of the error between the approximate and exact values of the integral over the sub-region (a, b) can be obtained as the difference between the more and less accurate approximations ($Q_1 = I_1$ and $Q_2 = I_2 + I_3$, respectively). This error is then compared to the local and global convergence criteria to determine if the error is acceptable (local and global convergence criteria are discussed in detail later in the chapter).

Step 4(i): If convergence is reached for the sub-region considered then the approximation of this particular sub-region is added to the accumulating approximation of the integral. Given that the order of the two approximations Q_1 and Q_2 are $O\left(h^4\right)$ and $O\left(\left[h/2\right]^4\right)$, respectively, Richardson extrapolation [23] can be used to combine these two estimates to obtain an integral estimate with an error of the order $O\left(\left[h/2\right]^8\right)$ by:

$$\Delta I = Q_2 + \frac{1}{15}\left(Q_2 - Q_1\right) \tag{7-9}$$

This high order estimate is the approximation that is added to the accumulating approximation of the integral. As convergence for this sub-region has been reached then the algorithm moves to the next sub-region (i.e. Step 1).



Step 1: Evaluate integrand at points a, b and c and use Simpsons rule to approximate the integral over region (a,b), I_1

Step 2: Evaluate integrand at points d and e and use Simpsons rule to approximate the integral over region (a,c), I_2, and over region (c,b), I_3.

Step 3: Compute error between approximations in steps 1 and 2 and compare error with local and global convergence criteria

Step 4(i): If convergence reached for sub-region considered then add increment to accumulating approximation of integral. Then obtain stored data for next sub-region (see step 4(ii)) and return to step 1.

Step 4(ii): If convergence is not reached then select the half of the current sub-region which has the larger integral approximation increment. Store the abscissa and function values for the region which is not selected (these values are retrieved at a later date), and then return to step 1 with the new sub-region.

Figure 7-3: Schematic illustration of the four steps in Magnitude-oriented Adaptive Quadrature (MAQ).

Step 4(ii): If convergence is not reached then the size of the sub-region needs to be reduced (and more function evaluations performed) so that the desired accuracy can be achieved. The sub-region is halved and the choice of which half to select is based on the estimates of the integral over the two half regions. For example, if $I_2 > I_3$ (i.e. the integral over region (a,c) is larger than over (c,b)) then region (a,c) is selected as the new sub-region. The abscissa and integrand values computed within the smaller sub-region are stored in an array and the algorithm returns to step 1 with the smaller sub-region. This smaller region is the region which will be the 'next sub-region' described in step 4(i) once convergence is reached in the larger sub-region.

For clarity in Figure 7-3, details on the monitoring the number of function evaluations performed relative to the maximum number prescribed, as well as details on storing of data points have been omitted. Further details can be found in the aforementioned programmed version of the algorithm which is available.

From the above four steps it becomes apparent that one of they key aims of the algorithm is to evaluate the integrand primarily around the region which contributes the most toward the integral (this is the effect in step 4(ii) of selecting the larger of the two half regions). This allows the initial computational effort to be focused on efficiently obtaining the dominant contribution to the integral. Figure 7-4 illustrates an example problem of the integration of the standard normal density function over the region x = (-5,5), with an error tolerance of tol = $1x10^{-3}$ using MAQ and adaptive quadrature (note that Romberg integration is not used in the comparison as it is not a locally adaptive method). It is noted for clarity that the standard normal density (SND) has the highest value about x = 0 with values tending to zero moving away from this maximum. As conventional adaptive quadrature always selects the left-hand sub-region (see step 4(ii) described previously) when convergence is not reached then adaptive quadrature effectively computes the integral from left-to-right. This is clearly illustrated in the accumulating value of the integral, where at the two tails of the SND the contribution to the integral is minor (represented by the initial and final ranges of function evaluations in Figure 7-4) and the dominant contribution occurs over a relatively small range of the region of integration (i.e. most of the contribution to the integral comes during a small window of function evaluations). In comparison, the MAQ algorithm targets the evaluation of the integral over this primary region first. As seen in Figure 7-4 after approximately 20 function evaluations, in which convergence is reached in three sub-regions, the accumulating value of the integral is already 40% of the final value (indicating most of the converged sub-regions were around x = 0).

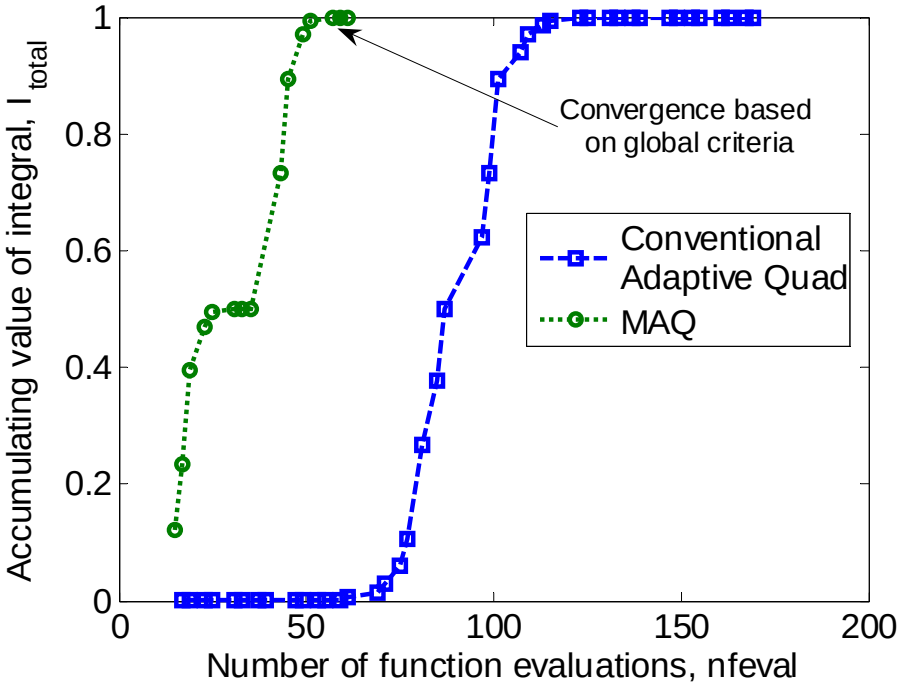


Figure 7-4: Example problem illustrating approximation of the standard normal density function integral as a function of number of function evaluations.

This targeting of the primary region of contribution to the integral is paramount in allowing the MAQ algorithm to also have global convergence criteria. This global convergence criterion is put in place to allow convergence over regions in which the contribution to the total integral is minor, despite local convergence not being reached. This is achieved by having two separate criteria for convergence (see Step 3 in Figure 7-3) only one of which is required to be satisfied. The first criterion is that of local convergence based on the truncation error in a single sub-region due to the step size between integration points. The second is that of global convergence which aims to allow convergence of a sub-region in which the integral contribution is minor compared to the final value of the integral. This is shown in Figure 7-4 for the MAQ algorithm by the high density of points which do not contribute significantly.

One problem encountered when constructing the MAQ algorithm is that: as the exact value of the integral is not known then how is it possible to determine if integration over some sub-region does or does not contribute significantly to the magnitude of the integral? As the current estimate of the integral is based on the integrand in regions where convergence has been achieved and the integrand is non-negative, this approximation will be less than the true value of the integral (i.e. global convergence of the integral is from below as shown in Figure 7-4), then the global tolerance criteria will be conservative. The conservatism is limited in the

sense that by the time the less significant regions of the integrand are approached, the majority of the integral has already been computed (and therefore the current accumulating approximation of the integral is likely close to the final value). Thus the only input requirements for the MAQ algorithm are (in general) two error tolerance values (one global and one local) and the maximum number of function evaluations. Herein we will drop the generalisation of using two different error tolerances and use the same error tolerance for both local and global convergence.

7.7 Efficiency comparison of integration methods

To illustrate the benefits of the proposed MAQ algorithm over Romberg integration and conventional adaptive quadrature, we investigate the two problems that were mentioned in the preceding sections.

7.7.1 Case 1: Expected annual loss computation

Firstly we investigate the case where the region of integration cannot be simply defined as is the case for computation of the expectation of the annual loss. We consider the same problem as referred to in Figure 7-2, and consider computation of the integral for relative error tolerances of 10^{-2} and 10^{-3}.

Figure 7-5 shows six subplots, which illustrate the distribution of the points of function evaluation for the three integration methods using the two different error tolerances. Here the phrase 'nfevals' is used to represent 'number of function evaluations', while 'tol' is used to represent the acceptable error tolerance. It can be seen that using Romberg Integration (Figures 7-5a and 7-5b) a significant number of function evaluations are expended computing the integrand for t values less than 0.5 (i.e. IM > 1.0). For both tolerance cases, convergence of the integral using Romberg Integration is primarily reduced due to a significant number of function evaluations being required to capture the two peaks of the integrand at t ~ 0.83 and 0.94.

Inspection of the results using adaptive quadrature (Figures 7-5a and 7-5d) illustrates that by measuring convergence locally, over regions of low curvature (such as t = 0.55 – 0.75, Figure 7-5c) the use of adaptive quadrature, as opposed to Romberg Integration can result in significant computational reductions. However, the flaw of measuring convergence only locally results in significant functions evaluations being required to evaluate the integrand for

$t < 0.5$, and around the region of $t = 1.0$, despite both of these regions contributing insignificantly toward the integrand.

Figures 7-5e and 7-5f illustrate the distribution of function evaluations required to compute the integral using MAQ. It can be seen that MAQ does not suffer from the two aforementioned problems exhibited by Romberg Integration and adaptive quadrature. Firstly, over the regions of high curvature ($t = 0.77$, 0.83, and 0.94) the algorithm adapts locally to reduce the step size around these regions. Secondly, over the insignificant regions of the integrand (i.e. $t < 0.5$), MAQ relaxes the local convergence requirements on the basis that the global contribution of this region to the value of the integral is less than the global tolerance of the current value of the integral. The reduction in function evaluations in the insignificant region is particularly evident when the error tolerance is reduced to 10^{-3} (Figure 7-5f). For an error tolerance of 10^{-2}, the use of MAQ results in a computational reduction of 4.5- and 8.8-fold compared to Romberg and Adaptive quadrature, respectively. For an error tolerance of 10^{-3}, the computational reductions are 7.9- and 6.8-fold, respectively.

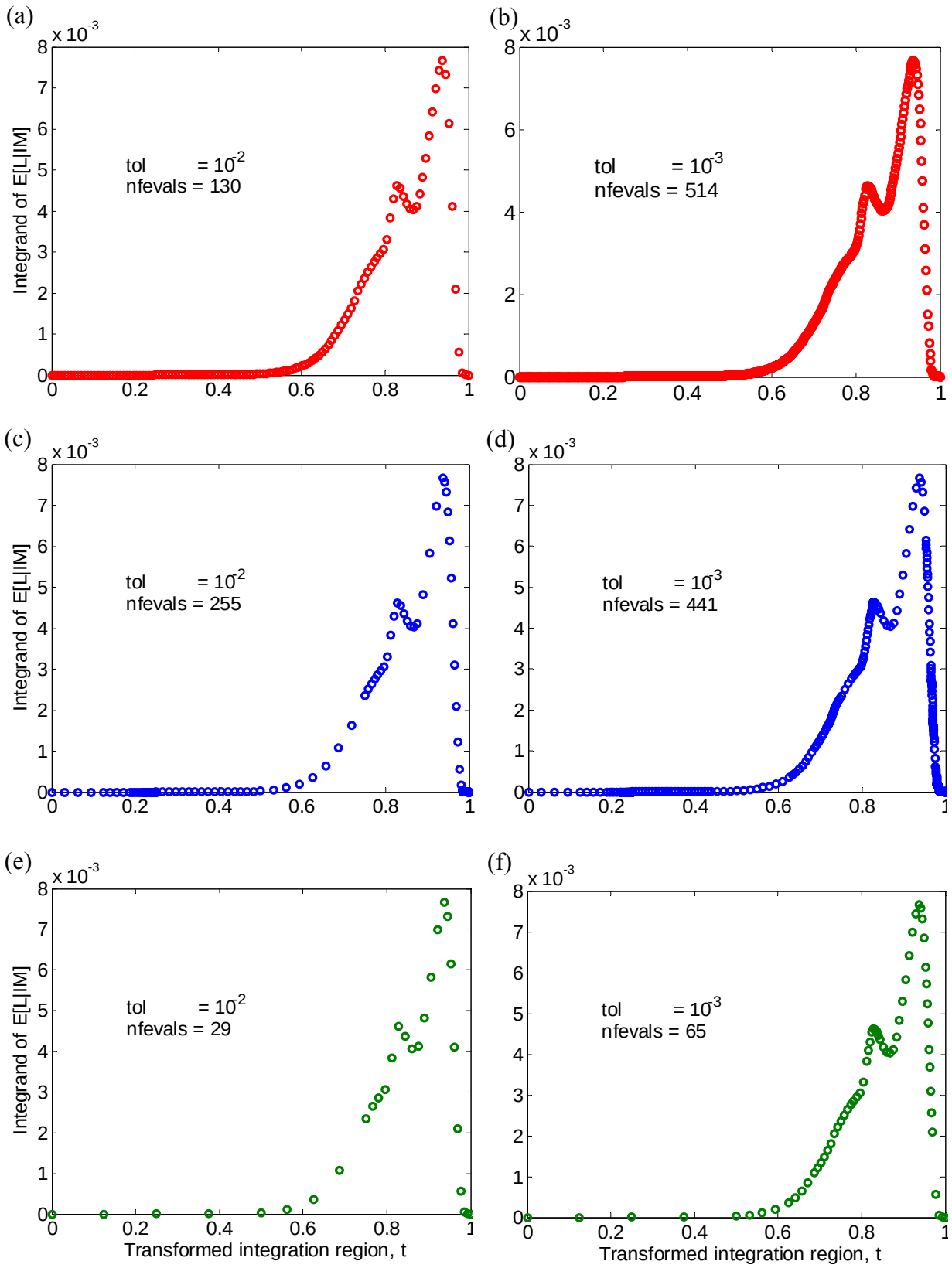


Figure 7-5: Illustration of the location of integrand evaluations for computing the expected annual loss problem for: (a)&(b) Romberg Integration; (c)&(d) Conventional adaptive quadrature; and (e)&(f) MAQ, for error tolerances of 10^{-2} and 10^{-3}.

7.7.2 Case 2: Probability of collapse computation

To illustrate the computation of a risk integral where one of the integrand arguments is a 'conventional' probability distribution we consider the computation of the annual probability of collapse. As previously mentioned, the fact that one of the arguments of the integrand is a probability distribution allows integration over a specified number of standard deviations, as opposed to integration over the entire domain of the integration variable. Similar to case 1 above, the efficiency (in terms of function evaluations) of the three integration methods is compared for error tolerances of 10^{-2} and 10^{-3}. Therefore the integral is computed based on using endpoints which are three standard deviations either side of the mean (which based on the standard normal distribution will give rise to an error of $2\text{x}10^{-3}$, due to neglecting the end regions).

We consider the ground motion hazard shown in Fig. 7-2b and assume that structure has a collapse fragility curve (assumed lognormal) with mean Spectral acceleration (Sa(T=2.0s,5%)) of 0.4g and dispersion of 0.3. As three standard deviations either side of the mean are considered, then the corresponding region of integration is IM = $\mu \exp\left(-0.5\beta^2 \pm 3\beta\right)$ = 0.085-1.714g, which corresponds to t = 0.368-0.922 (using Equation (7-5)).

Figure 7-6 illustrates the function evaluation distributions for the six different integration permutations. Again, similar to case 1, it can be seen that Romberg integration (Figures 7-6a and 7-6b) requires additional functional evaluations about the peak of the integrand to achieve global convergence, and as a result a significant number of function evaluations are expended over the tails of the integrand. Integration using adaptive quadrature (Figures 7-6c and 7-6d) captures the primary region of the integrand efficiently, but the local convergence criterion means that significant function evaluations are required over the tails of the distribution. Again, the use of MAQ for this problem allows efficient integration over the entire integration domain. For a tolerance of 10^{-2}, MAQ requires approximately only two-thirds of the function evaluations that are required using the other two methods. The efficiency of MAQ is particularly evident when the error tolerance is decreased to 10^{-3} (Figures 7-6e and 7-6f). Using MAQ, only 8 (~40%) additional function evaluations are required to achieve convergence for an error tolerance of 10^{-2}, while the number of function evaluations almost doubles for both Romberg Integration and Adaptive Quadrature.

It should also be noted although not shown here, that if a higher accuracy solution is

desired then the number of standard deviations to integrate over must be increased. This would result in substantially more computational demand for the Romberg and Adaptive quadrature algorithms, while the increase in computational demand for the MAQ algorithm would be only minor. The reason that the increase would be minor comes back to the idea that MAQ is efficient at handling these tails of the integrand (such as that shown Figure 7-4f for $t < 0.5$).

The MAQ algorithm can also be extended to the numerical evaluation of multiple integrals, in which it would be significantly more efficient compared to the other two alternatives discussed in this chapter since the computational work for multiple integrals can be approximated as $\overline{nfeval}^{N}$, where $\overline{nfeval}$ is the average number of function evaluations in each dimension, and N is the order/dimension of integration (i.e. single = 1, double = 2). Thus a two-fold reduction in computational work for a single integral is equivalent to a 4-fold reduction for a similar double integral.

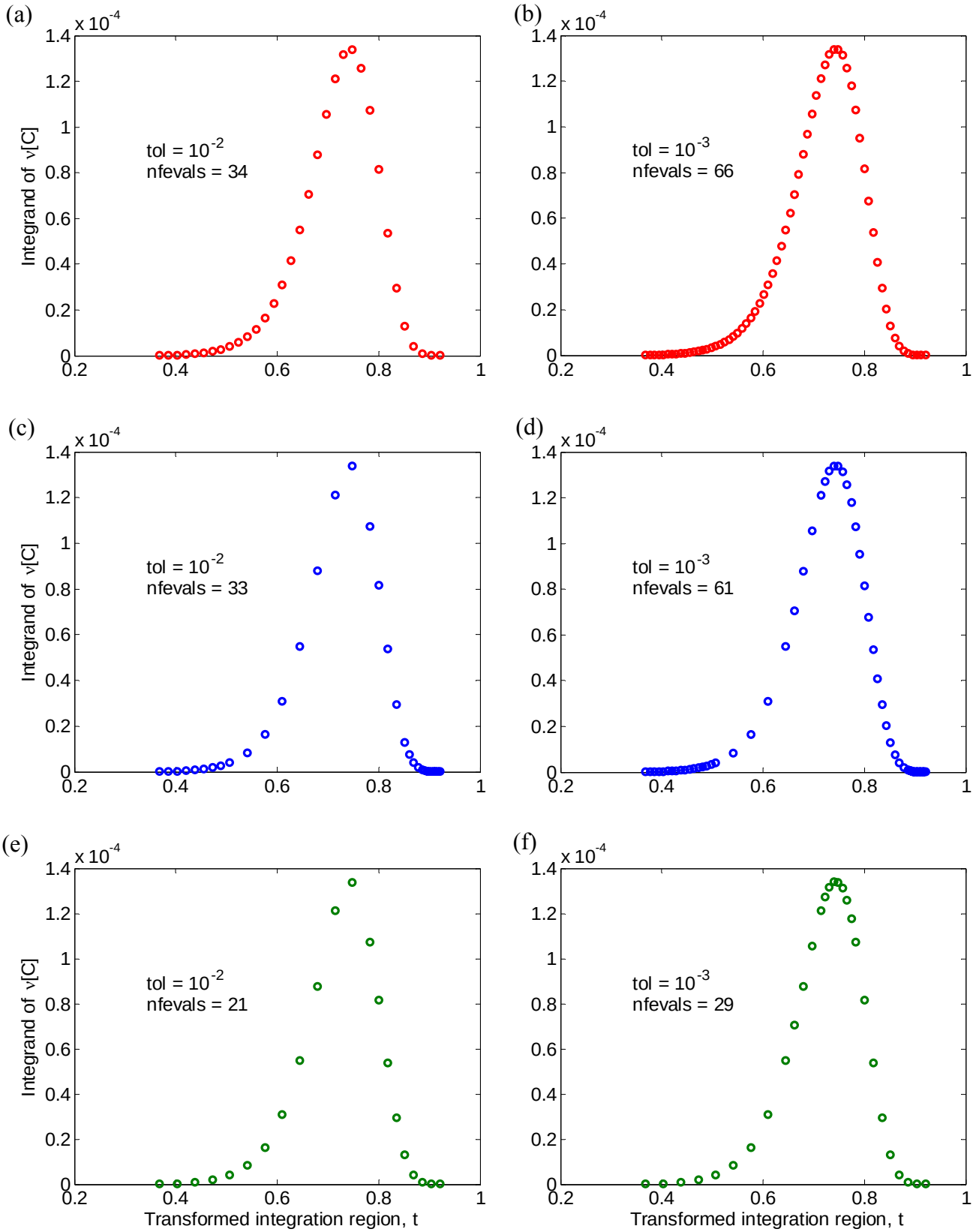


Figure 7-6: Illustration of the location of integrand evaluations for computing the annual rate of collapse problem using 5 standard deviations for: (a)&(b) Conventional adaptive quadrature; (c)&(d) Romberg Integration; and (e)&(f) MAQ, for error tolerances of 10^{-2} and 10^{-3}.

7.8 Conclusions

An integration algorithm called Magnitude-oriented Adaptive Quadrature (MAQ) was developed specifically for integration of the governing equations within the PEER framework formula. The MAQ algorithm allows integration over the entire integration domain, with computational effort focused on those regions which provide the majority of the contribution of the magnitude of the integral. It was illustrated that in addition to the MAQ algorithm being significantly more efficient than conventional integration algorithms such as Romberg integration and Simpson's-rule based Adaptive Quadrature, it is also 'user-friendly' in the sense that only the error tolerance is required to be specified (and not any information on the region of integration etc.).

Through two case examples it was shown that the computational reduction compared to Romberg and Adaptive quadrature algorithms can range from 4.5- to 8.8-fold in the case of integration over the entire integration domain; and a 30-100% reduction when only some portion of the integration region is required to be integrated. Therefore it can be conclusively stated that a seismic loss estimation framework which propagates uncertainty through direct numerical integration with MAQ will be significantly more efficient than one which uses conventional integration algorithms.

Although attention on the application of the proposed MAQ algorithm in this chapter has been given only to the PEER framing equations, there is no reason why the algorithm cannot be used in other situations. It will likely be an efficient integration algorithm for problems which contain integrals with a non-negative integrand of a similar functional form to those discussed here.

7.9 Acknowledgements

Financial support of the first author by the New Zealand Tertiary Education Commission is greatly appreciated.

7.10 References

[1] McGuire RK. Seismic Hazard and Risk Analysis. Earthquake Engineering Research Institute, 2004; 221.

[2] Cornell CA. Reliability-based earthquake-resistant design—the future, in 11th World Conference on Earthquake Engineering, Acapulco, Mexico, 1996.

[3] Kennedy RP and Short SA. Basis for Seismic Provisions of DOE-STD-1020. Lawrence Livermore National Laboratory and Brookhaven National Laboratory, Washington, D.C., 1994.

[4] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[5] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1901-1914.

[6] Porter KA, Kiremidjian AS, and LeGrue JS. Assembly-Based Vulnerability for Buildings and Its Use in Performance Evaluation. Earthquake Spectra 2001; 17(2): 291-312.

[7] Au SK and Beck JL. Subset Simulation and its Application to Seismic Risk based on Dynamic Analysis. Journal of Engineering Mechanics 2003; 129(8): 901–917.

[8] Deierlein GG, Krawinkler H, and Cornell CA. A framework for performance-based earthquake engineering, in 7th Pacific Conference on Earthquake Engineering, Christchurch, New Zealand, 2003.

[9] Evans M and Swartz T. Methods for Approximating Integrals in Statistics with Special Emphasis on Bayesian Integration Problems. Statistical Science 1995; 10(3): 254-272.

[10] Baker JW and Cornell CA. Uncertainty propagation in probabilistic seismic loss estimation. Structural Safety 2008; 30(3): 236-252.

[11] Bradley BA and Lee DS. Accuracy of approximate methods of uncertainty propagation in loss estimation. Structural Safety 2009; (in press).

[12] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[13] Baker JW. Vector-valued Ground Motion Intensity Measures for Probabilistic Seismic Demand Analysis. Ph.D Thesis, Department of Civil and Environmental Engineering Stanford University, 2005, 347.

[14] Bazzurro P. Probabilistic seismic demand analysis. Stanford University, Stanford, CA, 1998. 329. http://www.stanford.edu/group/rms/

[15] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[16] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges. Engineering Structures 2007; 29(10): 2662-2672, DOI: 10.1016/j.engstrut.2006.12.015.

[17] Aslani H and Miranda E. Probability-based Seismic Response Analysis. Engineering Structures 2005; 27(8): 1151-1163.

[18] Luco N, Cornell CA, and Yeo GL. Annual limit state frequencies for partially-inspected earthquake-damage buildings. Structural Safety 2002; 24(2-4): 281-296.

[19] Benjamin JR and Cornell CA. Probability, statistics, and decision for civil engineers. McGraw-Hill: New York, 1970;

[20] Bradley BA, Dhakal RP, Cubrinovski M, Mander JB, and MacRae GA. Improved seismic hazard model with application to probabilistic seismic demand analysis. Earthquake Engineering and Structural Dynamics 2007; 36(14): 2211-2225.

[21] Bradley BA, Dhakal RP, and Mander JB. Parametric structure specific seismic loss estimation. 4th International Conference on Urban Earthquake Engineering, Tokyo, Japan, 2007.

[22] Kythe PK and Schaferhotter MR. Handbook of Computational Methods for Integration. Chapman and Hall/CRC, 2000;

[23] Chapra SC and Canale RP. Numerical Methods for Engineers: with software and programming applications. Fourth Edition ed. McGraw-Hill, 2002;

8. Accuracy of Approximate Methods of Uncertainty Propagation in Seismic Loss Estimation

Bradley BA, Lee DS. Accuracy of approximate methods of uncertainty propagation in seismic loss estimation. *Structural Safety* 2009. (in press).

8.1 Abstract

In this chapter the efficacy of an approximate method of uncertainty propagation, known as the first-order second-moment (FOSM) method, for use in seismic loss estimation is investigated. The governing probabilistic equations which define the Pacific Earthquake Engineering Research (PEER)-based loss estimation methodology used are discussed, and the proposed locations to use the FOSM approximations identified. The justification for the use of these approximations is based on a significant reduction in computational time by not requiring direct numerical integration, and the fact that only the first two moments of the distribution are known. Via various examples it is shown that great care should be taken in the use of such approximations, particularly considering the large uncertainties that must be propagated in a seismic loss assessment. Finally, a complete loss assessment of a structure is considered to investigate in detail the location where significant approximation errors are incurred, where caution must be taken in the interpretation of the results, and the computational demand of the various alternatives.

8.2 Introduction

There are many uncertainties in earthquake engineering, ranging from uncertainties in the spatial and temporal occurrence and size distribution of earthquakes, to the uncertainties in the structural response for a given level of ground motion shaking, and the damage and loss consequences as a result of the structural response. In contemporary performance-based earthquake engineering (PBEE), there is a need to account for such uncertainties when characterising the performance of a structural system vulnerable to earthquake hazards. Great progress has been made in frameworks which consistently account for all of the aforementioned uncertainties, primarily since the presentation of the so-called "PEER framing formula" [1, 2] advocated by the Pacific Earthquake Engineering Research (PEER) centre. The PEER framing formula allows explicit and consistent treatment of uncertainties in the process of PBEE. The key assumption of conditional independence allows the framing formula to be separated into four different relationships, which are typically solved by different expert personnel, and then integrated together to provide decision variables for stakeholders to make rational decisions regarding seismic risk.

In order to conduct loss assessments within such a framework by combining the aforementioned relationships, the uncertainties in each of the steps must be propagated through to the final decision variables. This uncertainty propagation can be done in various ways with various levels of sophistication and accuracy. The primarily continuous nature of the relationships which comprise the PEER framing formula results in uncertainty propagation via integration over the domain of the random variables. Despite the ever increasing processing abilities of modern computing, as the dimensionality of the integral equations increase, the computational work to perform seismic loss estimation can still be demanding. As a result, there is still a desire by many to develop uncertainty propagation methods which allow results to be obtained efficiently, without significant loss of accuracy. In particular, the moment-method approach of approximating probability density functions instead of using approximate (i.e. numerical) integration methods has become popular. In particular, second-moment methods, in which only the first two moments (mean and standard deviation) of the random variables are retained, have been used widely in design code calibration [e.g. 3].

Regarding uncertainty propagation specifically with the PEER seismic loss estimation framework, use of the first-order second-moment (FOSM) method has been proposed by Baker and Cornell [4, 5]. The FOSM method uses a first-order Taylor Series approximation

of a function of random variables, with each random variable being approximated by its first two moments. Baker and Cornell use the FOSM method primarily for computation of the relationship between a vector of losses and ground motion intensity, which they identify as the computationally demanding step of the seismic loss estimation process. A similar approach has also been taken by Aslani [6], who uses the FOSM method for determination of the covariance structure of the loss given intensity relationship. Despite this use of the FOSM method in seismic loss estimation, as far as the authors are aware, to date its accuracy and efficiency in reducing the computational demand has not yet been scrutinized in detail.

In this chapter, the use of the FOSM method in various stages of the PEER framework formula is investigated. Its limitations regarding accuracy in the computation of various measures of seismic performance is illustrated through application to the loss assessment of a building, and comparison with the direct solution via numerical integration. The computational demands of the FOSM methods are also discussed in order for the benefits and shortfalls of the method to be properly put in perspective.

8.3 Probabilistic seismic loss estimation framework

For completeness, a review of the PEER loss assessment framework equations is given in this section, which also forms the basis for the FOSM approximations investigated throughout the remainder of this chapter. In this section we consider aleatory uncertainties only. The inclusion and propagation of epistemic uncertainties is discussed briefly later in the chapter.

The PEER loss estimation framework is built around the following equation [5]:

$$\lambda_{L_T}(z) = \int_{\mathbf{u}}\int_{\mathbf{v}}\int_{\mathbf{y}}\int_{x} G_{L_T|\mathbf{L}}(z\,|\,\mathbf{u}) f_{\mathbf{L}|\mathbf{DS}}(\mathbf{u}\,|\,\mathbf{v}) f_{\mathbf{DS}|\mathbf{EDP}}(\mathbf{v}\,|\,\mathbf{y})\, f_{\mathbf{EDP}|IM}(\mathbf{y}\,|\,x)\left|\frac{\mathrm{d}\lambda_{IM}(x)}{\mathrm{d}x}\right| \mathbf{d}\mathbf{u}\mathbf{d}\mathbf{v}\mathbf{d}\mathbf{y}\mathrm{d}x \qquad (8\text{-}1)$$

where IM is an intensity measure representing the intensity of the ground motion observed at the site of the structure; $\boldsymbol{EDP}|IM$ is a vector of engineering demand parameters which characterise the response of the structure for a given level of ground motion $IM = im$; $\boldsymbol{DS}|\boldsymbol{EDP}$ is a vector of damage states for each component given demand $\boldsymbol{EDP} = \boldsymbol{edp}$; $\boldsymbol{L}|\boldsymbol{DS}$ is a vector of losses incurred in each component due to damage states, $\boldsymbol{DS} = ds$; L_T is the total loss in the structure; $G_{X|Y}(x|y) = \mathrm{P}(X \geq x | Y = y)$ is the complementary cumulative distribution function (CCDF) of X given $Y = y$; $f_{X|Y}(x|y)$ is the probability density of X given $Y = y$; and $\lambda_Z(z)$ is the annual frequency of Z exceeding z. Herein for brevity we will use the shorthand notation $G_{X|Y}$

, $f_{X|Y}$, and λ_Z in place of $G_{X|Y}(x|y)$, $f_{X|Y}(x|y)$, and $\lambda_Z(z)$, respectively. Equation (8-1) represents an application of the total probability theorem. As previously mentioned, Equation (8-1) also makes the conditional independence assumption (e.g. that the damage state, ***DS***, is only a function of ***EDP***, and not of the *IM* causing ***EDP***), and therefore can be de-coupled and solved in separate stages.

All components which are used to describe the structure and its inventory are defined by fragility and loss functions. Fragility and loss functions can be combined (for a single component) to obtain the relationship between component loss for a given *EDP* (*L*|*EDP*) (e.g. [6]). From the results of seismic response analysis a relationship can be found between ground motion intensity measures (IM's) and various *EDP*'s throughout the structure. This *EDP*|*IM* relationship can be combined with the *L*|*EDP* relationship to obtain a relation between the loss for a single component, *k*, (which is dependent on EDP_i) given a certain level of *IM* (L_k|*IM*):

$$\mu_{L_k|IM}(im) = \int \mu_{L_k|EDP_i}(edp) f_{EDP_i|IM}(edp|im) \mathrm{d}EDP \tag{8-2}$$

$$\sigma^2_{L_k|IM}(im) = \int \left[\mu^2_{L_k|EDP_i}(edp) + \sigma^2_{L_k|EDP_i}(edp)\right] f_{EDP_i|IM}(edp|im) \mathrm{d}EDP - \mu^2_{L_k|IM}(im) \tag{8-3}$$

where $\mu_{L_k|IM}$ is the mean loss for component *k* given $IM = im$; $f_{EDP_i|IM}$ is the probability density function of EDP_i given $IM = im$; $\mu_{L_k|EDP_i}$ is the mean loss for component *k* given $EDP_i = edp_i$; and $\sigma^2_{L_k|EDP_i}$ = the variance in the loss for component *k* given $EDP_i = edp_i$.

Equations (8-2) and (8-3) can be used to obtain the first two moments for the distribution of loss given *IM* for a single component *k*. For the case where the structure does not collapse it is reasonably assumed that the total direct repair loss for the entire structure can be obtained from the summation of the direct repair losses in each of the different components. More specifically:

$$\mu_{L_T|IM,NC}(im) = \sum_{k=1}^{Nc} \mu_{L_k|IM}(im) \tag{8-4}$$

$$\sigma^2_{L_T|IM,NC}(im) = \sum_{k=1}^{Nc} \sigma^2_{L_k|IM}(im) + 2\sum_{k=1}^{Nc}\sum_{k'=1}^{k-1} \sigma_{L_k|IM,L_{k'}|IM}(im) \tag{8-5}$$

where N_c is the number of components in the structure; and $\sigma_{L_k|IM,L_{k'}|IM}$ is the covariance in the *L*|*IM* relationship between components *k* and *k'*. This covariance can be computed by:

$$\sigma_{L_k|IM,L_{k'}|IM}(im) = \mu_{L_kL_{k'}|IM}(im) - \mu_{L_k|IM}(im)\mu_{L_{k'}|IM}(im) \tag{8-6}$$

In which $\mu_{L_kL_{k'}|IM}(im)$ is computed from:

$$\mu_{L_kL_{k'}|IM}(im) = \int\int \left(\mu_{L_k|EDP_i}(edp_i)\mu_{L_{k'}|EDP_{i'}}(edp_{i'}) + \sigma_{L_k|EDP_i,L_{k'}|EDP_{i'}}(edp_i, edp_{i'})\right) \text{x} \\ f_{EDP_i,EDP_{i'}|IM}(edp_i, edp_{i'}|im)\text{d}EDP_i\text{d}EDP_{i'} \tag{8-7}$$

where $f_{EDP_i,EDP_{i'}|IM}(edp_i, edp_{i'}|im)$ is the joint distribution of EDP_i and $EDP_{i'}$ given $IM = im$; and $\sigma_{L_k|EDP_i,L_{k'}|EDP_{i'}}(edp_i, edp_{i'})$ is the covariance in the $L|EDP$ relationship for components k and k'.

By considering the mutually exclusive and collective exhaustive events of collapse and no collapse, the expected loss for a given IM can be computed from:

$$\mu_{L_T|IM}(im) = \mu_{L_T|IM,NC}(im)\left[1 - P_{C|IM}(im)\right] + \mu_{L_T|C}P_{C|IM}(im) \tag{8-8}$$

where $\mu_{L_T|C}$ is the mean of the loss given global collapse (assumed independent of IM); $\mu_{L_T|IM}(im)$ is the mean of the total loss once conditioning on collapse (and no collapse) has been removed; and $P_{C|IM}(im)$ is the probability of collapse given $IM = im$. The standard deviation of the total loss given IM, $\sigma^2_{L_T|IM}(im)$, although not shown here can be computed in a similar manner (See [5, 7] for details).

The expected annual loss can be computed by integrating the expected loss for a given level of ground motion intensity, $\mu_{L_T|IM}(im)$, with the ground motion hazard curve, $\lambda_{IM}(im)$, which gives the annual frequency of exceeding different levels of ground motion intensity.

$$\mu_{L_T} = \int \mu_{L_T|IM}(im)\left|\frac{d\lambda_{IM}(im)}{dIM}\right|dIM \tag{8-9}$$

The loss hazard curve gives the annual frequency of exceeding a specified value of loss, and can be computed from:

$$\lambda_{L_T}(l_t) = \int G_{L_T|IM}(l_t|im)\left|\frac{d\lambda_{IM}(im)}{dIM}\right|dIM \tag{8-10}$$

where $G_{L_T|IM}(l_t|im)$ is the CCDF of the total loss given $IM = im$.

8.4 Approximate uncertainty propagation

It is proposed herein to investigate the efficacy of using the FOSM method in place of some of the equations in the integral-based loss estimation methodology presented in the previous section. Initially, the same stance as that of Baker and Cornell [5] is taken for the location of the FOSM approximations. That is, the FOSM method is used primarily in computing the moments of the *L|IM* relationship since in most practical situations there will be many different components and therefore this step of the loss estimation methodology is potentially computationally intensive. Direct numerical integration is used for equations with the ground motion hazard (i.e. Equations (8-9) and (8-10)), since the ground motion hazard is a dominant contributor toward the total uncertainty, and also the dimensionality of the integral is one (hence the computational demand is relatively small in comparison to the computation of the moments of the *L|IM* relation). The following section gives an overview of the development of the FOSM method and its limitations to some general problems; its use for the particular seismic loss assessment problem is discussed in subsequent sections.

8.4.1 First order second moment (FOSM) method

Consider a general scalar function $Y = g(X)$ where X is a random variable. The moments (mean, variance, etc) of Y can be obtained directly by integrating the product of $g(X)$ and $f_X(x)$ over the support of X [8] (the particular form of the product depends on the moment being computed), where $f_X(x)$ is the probability density function of X. Under certain conditions, it may be appropriate to approximate these moments based on a Taylor series expansion of $g(X)$ about the mean value of x, that is:

$$g(X) = g(\mu_X) + (X - \mu_X)\frac{dg}{dX} + \frac{1}{2!}(X - \mu_X)^2 \frac{d^2 g}{dX^2} + ... \tag{8-11}$$

where all of the derivatives are evaluated at $X = \mu_X$.

From Equation (8-11) and using the fact that $E[Y] = E[g(X)]$ and $Var[Y] = E[g(X)^2] - E[g(X)]^2$, a first order estimate of the mean and variance of Y are:

$$E[Y] = \mu_Y \approx g(\mu_X) \tag{8-12}$$

$$Var[Y] = \sigma_Y^2 \approx \sigma_{g(X)}^2\Big|_{X=\mu_X} + \sigma_X^2 \left(\frac{dg}{dX}\right)^2\Bigg|_{X=\mu_X} \tag{8-13}$$

where $\sigma^2_{g(X)}$ is the variance in $g(X)$ and σ^2_X is the variance in X itself, both of which are evaluated at $X = \mu_X$. The derivative, $\frac{dg}{dX}$, can be thought of as the sensitivity of changes in X to the resulting Y. Equations (8-12) and (8-13) define what is known as the first-order-second moment (FOSM) method in structural reliability [9]. The FOSM method is useful as the first two moments of the function $Y = g(X)$ can be obtained without the need for integration, and only the first two moments of X are required. In some instances the second order estimate of the mean can also be used (which still requires only the first two moments of X). The expectation of the second order truncation of Equation (8-12) yields:

$$E[Y] = \mu_Y \approx g(\mu_X) + \frac{1}{2}\sigma_X^2 \left.\frac{d^2 g}{dX^2}\right|_{X=\mu_X} \tag{8-14}$$

Thus, unlike the first order estimate of the mean (Equation (8-12)) which is independent of the variance of X, the second order estimate of the mean (Equation (8-14)) accounts for the variance in X, which also depends on the second derivative (curvature) of $g(X)$. Below, two problems are considered to illustrate some salient features of the approximate methods given by Equations (8-12)- (8-14).

8.4.2 Example 1: $Y=X^2$.

In this first example consider the case $Y = g(X) = X^2$. It is assumed that X has a lognormal distribution with a mean, $\mu_X = 2$ and coefficient of variation (*COV*), $\delta_X = \sigma_X / \mu_X$ varying from 0.01 to 1.0. As $g(X)$ is quadratic (and thus its second derivative is constant) then the second order estimate of the mean (Equation (8-14)) is exact and forms the basis for the accuracy of the FOSM method in this case. The exact and FOSM approximations for the moments of Y are:

$$\mu_Y = {\mu_X}^2 + {\sigma_X}^2 \approx {\mu_X}^2 \tag{8-15}$$

$$\sigma_Y^2 = \left(\mu_X^2 + \sigma_X^2\right)^2 \left[\left(\frac{\sigma_X^2}{\mu_X^2} + 1\right)^4 - 1\right] \approx 4\sigma_X^2 \mu_X^2 \tag{8-16}$$

Figure 8-1 illustrates the corresponding errors in the first order mean and standard deviation estimations of $Y = X^2$. It is observed that the first order estimate of the mean always under predicts the value of μ_Y (as is obvious from Equation (8-15)), with the relative error increasing as the COV increases. The error in the standard deviation also increases as the

COV increases, and in particular, for a given COV the error in the standard deviation is notably higher than for the mean approximation. For example, for a COV of 0.6 the relative error in the mean is approximately -26%, while the error in the standard deviation is approximately -77%. These correspond to error ratios (i.e. the exact value divided by the FOSM value) of 1.35 and 3.1, respectively.

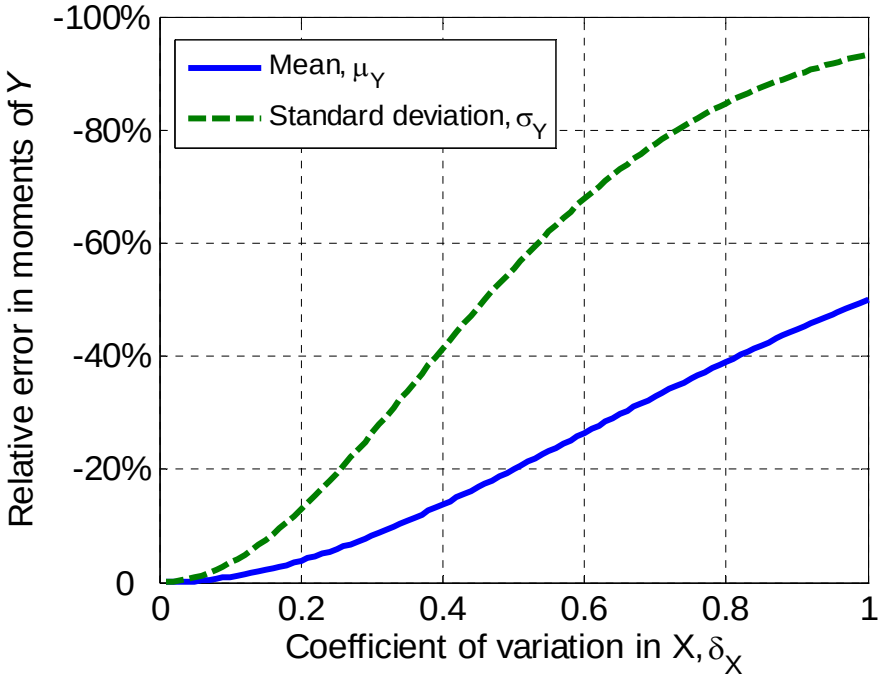


Figure 8-1: Errors in the mean and standard approximation for Example 1.

8.4.3 Example 2: Second-order mean approximation.

In the second example consider the use of both the first- and second-order methods for estimation of the expectation of $Y = g(X)$ for various levels of uncertainty in X. As the particular results may be dependent on the analytical form of $g(X)$, relationships are used which represent the general trends in the relationships between loss and demand ($L|EDP$), and between demand and intensity ($EDP|IM$) for the seismic loss estimation problem. The (normalised) expected $L|EDP$ relationship is based on a component with a single damage state having a lognormal fragility curve with mean, $\mu_{\ln EDP|DS} = -1.20$ (equivalent to a median of $\overline{EDP} = 0.3$) and lognormal standard deviation (dispersion), $\sigma_{\ln EDP|DS} = 0.3$. The expected $EDP|IM$ relationship used is of a power model form, $\mu_{EDP|IM} = 0.0025IM^2$, which has been used by numerous researchers.

Figure 8-2 illustrates the accuracy of the first- and second-order approximations for

various dispersion values in the *EDP*|*IM* relationship. Figure 8-2a illustrates that for small values of the dispersion ($\sigma_{lnEDP|IM} = 0.15$) the second-order approximation is significantly better than the first-order approximation, effectively lying over the exact curve obtained via numerical integration. For a moderate value of dispersion ($\sigma_{lnEDP|IM} = 0.25$), the accuracy of the second-order approximation is good for *IM* values which correspond to loss values below 0.5, but for loss values above 0.5, there is some deviation of the second-order approximation from the exact solution for larger *IM* values. Note that the first order approximation (Equation (8-12)) is independent of the uncertainty in the *EDP*|*IM* relationship, thus yielding the same result for all dispersion values.

Figure 8-2b illustrates the accuracy of the approximations for larger values of the dispersion in the *EDP*|*IM* relationship. In this case we can see that in the vicinity of $\mu_{L|IM} = 0.5$ there is significant error in the second-order approximation, which increases with increasing dispersion. Careful inspection of Equation (8-14) and the second derivative of the *L*|*EDP* relationship illustrates that this error is caused by the second derivative of the *L*|*EDP* relationship, which is largest either side of $\mu_{L|IM} = 0.5$. As it is by no means uncommon to have uncertainties this large in the *EDP*|*IM* relationship, then primarily because of this apparent 'instability' of the second-order method, we will use the first-order approximation of the mean throughout the remainder of this chapter.

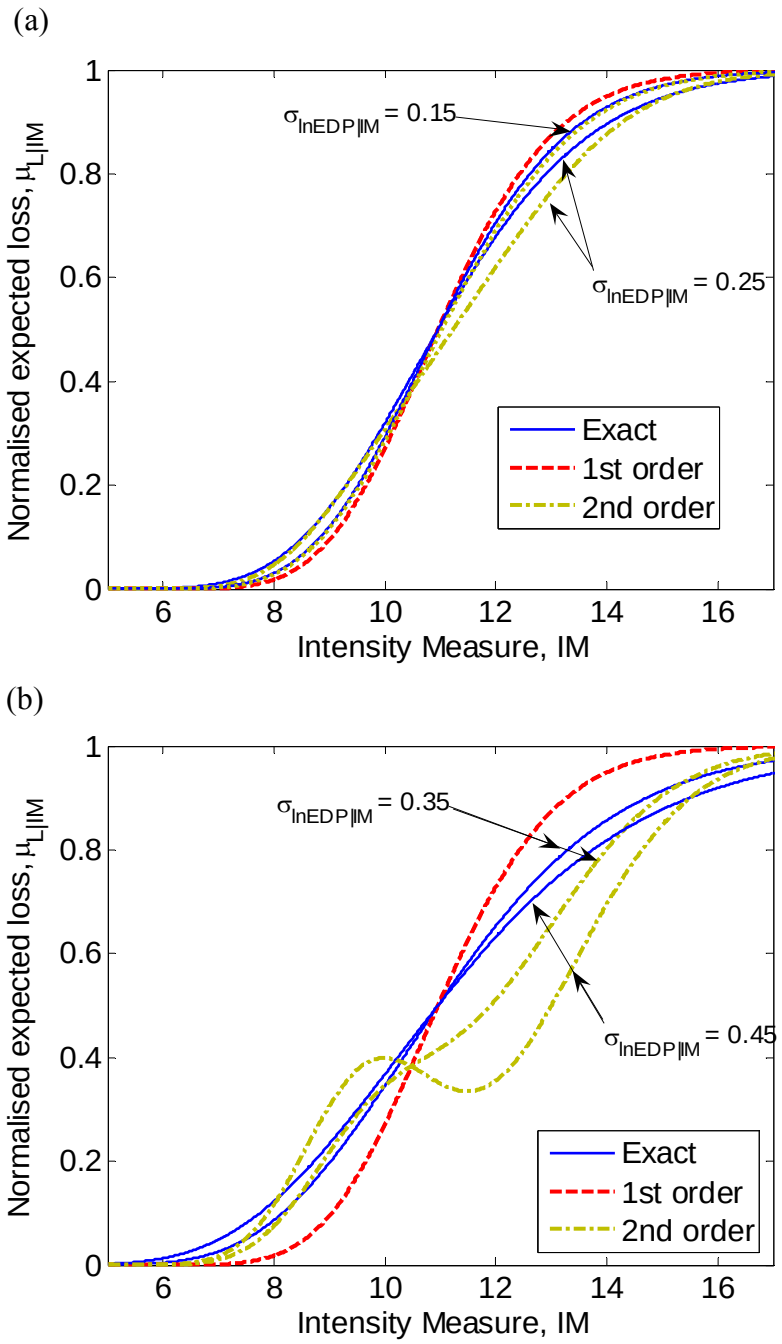


Figure 8-2: Accuracy of first- and second-order approximations in Example 2 for various demand-intensity dispersions, $\sigma_{\ln EDP|IM}$.

8.5 Approximate seismic loss assessment

In this section consideration is given to the use of the FOSM method within the seismic loss assessment framework. As it is common within the loss assessment framework to assume that the distribution of the demand given intensity relationship (*EDP*|*IM*) is lognormal then log-forms are used within the FOSM method (such that the ln*EDP*|*IM* distribution is normal)

and then are later converted back to non-log forms.

8.5.1 Expected loss given intensity measure, *E*[*L*|*IM*]

The FOSM approximation for the expected loss given *IM* for component *k* (computed directly using Equation (8-2)) can be shown to equal [5]:

$$\mu_{\ln L_k|IM} \cong \mu_{\ln L_k|\ln EDP_i}\left(\mu_{\ln EDP_i|IM}\right) \tag{8-17}$$

That is, $\mu_{\ln L_k|IM}$ is obtained by computing $\mu_{\ln L|\ln EDP_i}$ at the mean value of $\ln EDP_i|IM$, $\mu_{\ln EDP_i|IM}$. The non-log form is then obtained by the first-order approximation $\mu_{L_k|IM} \cong e^{\mu_{\ln L_k|IM}}$.

To investigate the accuracy of the FOSM approximation for $\mu_{L_k|IM}$, two different types of components are selected. The first is a reinforced concrete (RC) column and the second an interior partition. The properties of the damage state fragility and loss functions are given in Table 8-1. In Table 8-1, the *EDP* of peak interstorey drift (θ_{max}) values have been obtained from the deformation damage index (DDI) values given in Mitrani-Reiser [10] using recoverable elastic and ultimate rotations of $\theta_r = 0.005$ and $\theta_u = 0.06$, respectively. The repair costs for the RC column are per column costs, while for the interior partition the costs are per 64 ft^2 [10]. These two different components have been selected since the RC column has fragility curves with large uncertainty ($\sigma_{\ln EDP|DS}$ = 0.74 – 1.36), while the uncertainty in the partition fragility curves is approximately 25% of those of the column ($\sigma_{\ln EDP|DS}$ = 0.17, 0.23). The uncertainty in the fragility curves directly affects the second and higher order derivatives of the loss given demand relationship, and it will be shown that this affects the error in the FOSM approximation.

Table 8-1: Damage state and loss properties for the two components considered.

	RC column[1,2] [10]				Internal Partition[1] [10]	
	DS_1	DS_2	DS_3	DS_4	DS_1	DS_2
$\mu_{EDP\|DS}$	0.0044	0.017	0.039	0.070	0.0039	0.0085
$\sigma_{\ln EDP\|DS}$	1.36	0.89	0.80	0.74	0.17	0.23
$\mu_{L\|DS}$	8.0	22.5	34.3	34.3	0.088	0.525
$\sigma_{\ln L\|DS}$	0.42	0.40	0.37	0.37	0.2	0.2

[1]*EDP* is max interstory drift, and loss is in thousand USD.
[2]DDI values from Mitrani-Reiser [10] converted to θ_{max} using $\theta_r = 0.005$ and $\theta_u = 0.06$.

Figure 8-3a and Figure 8-3b illustrate the results for the expected loss given *IM* for the column due to DS_3 occurrence, for dispersions of 0.3 and 0.5 (which are assumed independent of *IM*) in the *EDP*|*IM* relationship. The mean *EDP*|*IM* relationship used corresponds to that of the first floor of the case study structure discussed in the final section of this chapter. It can be seen that as expected the error is directly related to the magnitude of $\sigma_{\ln EDP|IM}$, with a relatively small error when $\sigma_{\ln EDP|IM} = 0.3$ compared to the error when $\sigma_{\ln EDP|IM} = 0.5$. As in the previous section it is again noted that the FOSM method underestimates the 'exact' loss for small *IM* values and over estimates for large *IM* values. Figure 8-3c and Figure 8-3d illustrate the *L*|*IM* curves for the partition due to the onset of DS_2. Similar trends are seen with that of the column case, in regard to the error being proportional to $\sigma_{\ln EDP|IM}$, and also systematic underestimation and overestimation of the error for small and large *IM* values, respectively. It is also apparent that the error in the FOSM approximation for the partition is much larger than for the column component. Recall that the accuracy of the approximate methods depends not only on the uncertainty in *EDP*|*IM* relationship (which is the same for both the RC column and the interior partition), but also 'shape' (specifically the derivatives) of the *L*|*EDP* relationship (i.e. the second-order approximation (Equation (8-14)) indicates the mean depends on the second derivative of *L*|*EDP*). Thus, the smaller uncertainty in the partition component damage state, results in a *L*|*EDP* relationship which has significantly larger higher order derivatives around the *IM* causing a normalised loss of 0.5, relative to the RC column, and thus why the error is larger for the interior partition.

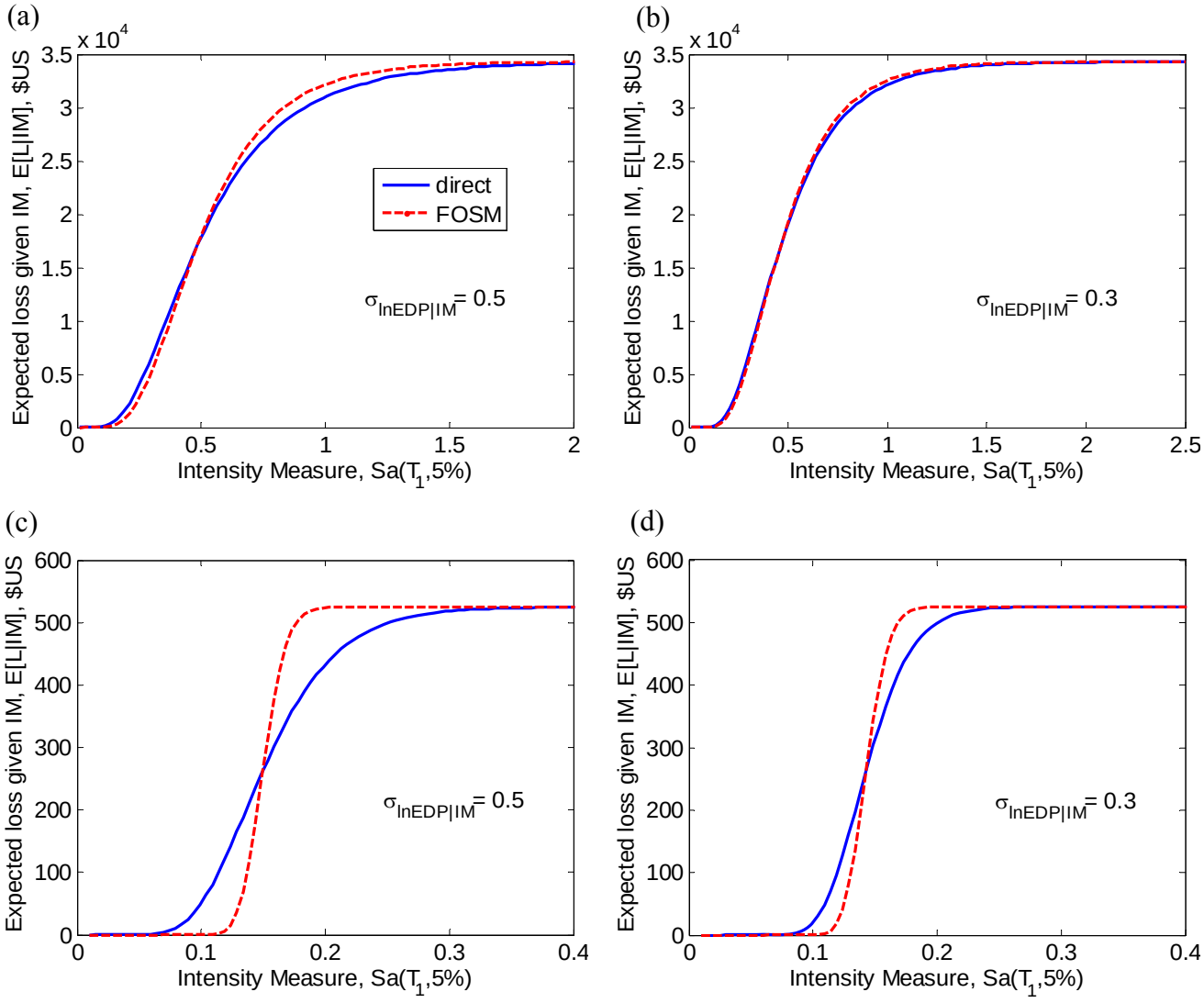


Figure 8-3: Accuracy of FOSM method for expected loss given *IM* computation for: (a)&(b) occurrence of DS_3 in the RC column component; and (c)&(d) occurrence of DS_2 in the partition component.

8.5.2 Standard deviation in loss given intensity measure, $\sigma[L|IM]$

Similar to the mean loss given intensity measure, the standard deviation in the loss given *IM* can be obtained using a FOSM approximation, which is given by [5]:

$$\sigma^2_{\ln L_k|IM} \cong \sigma^2_{\ln EDP_i|IM}\left(\frac{\partial \ln L_k|\ln EDP_i}{\partial \ln EDP_i}\right)^2\Bigg|_{\mu_{\ln EDP_i|IM}} + \sigma^2_{\ln L_k|\ln EDP_i}\Big|_{\mu_{\ln EDP_i|IM}} \tag{8-18}$$

Figure 8-4a and Figure 8-4b illustrate the dispersion, $\sigma_{\ln L_k|IM}$, in the $L|IM$ relationship for the column component due to the onset of DS_3 (i.e. analogous to the results in Figure 8-3a and Figure 8-3b for the expectation). Again, as for the mean $L|IM$ case, it is evident that the

error between the FOSM approximation and the direct method is a function of the dispersion in the *EDP*|*IM* relationship, $\sigma_{\ln EDP|IM}$. In both cases, it is also apparent that the FOSM approximation of $\sigma_{\ln L_k|IM}$ is larger than that obtained by direct integration. The dispersion must next be converted back to the (non-log) standard deviation, which can be achieved in several ways. Firstly, one can make the assumption that the *L*|*IM* relation has a lognormal distribution (this is the assumption made to get the dispersion from the computed standard deviation in the case of direct integration). Based on this assumption the standard deviation of *L*|*IM* can be obtained from [8]:

$$\sigma_{L_k|IM} = \mu_{L_k|IM}\sqrt{e^{\sigma^2_{\ln L_k|IM}} - 1} \tag{8-19}$$

The second approach is to use a first-order approximation [5] to obtain:

$$\sigma_{L_k|IM} \cong \mu_{L_k|IM}\sigma_{\ln L_k|IM} \tag{8-20}$$

Note that Equation (8-20) can be obtained from Equation (8-19) by using the Taylor Series approximation: $e^{\sigma^2} = 1 + \sigma^2 + O\left(\sigma^4\right)$.

Figure 8-4c and Figure 8-4d illustrate the comparison between the FOSM and direct approaches for computing $\sigma_{L_k|IM}$, using both Equations (8-19) and (8-20) above. It is evident that the first-order assumption (Equation (8-20)) causes an under approximation in $\sigma_{L_k|IM}$ even when the dispersion is well predicted (i.e. Figure 8-4a and Figure 8-4b). This occurs because the magnitude of the dispersion, $\sigma_{\ln L_k|IM}$, is large enough such that the higher order terms in the TS approximation for $e^{\sigma^2_{\ln L_k|IM}}$ are significant, particularly at small *IM* values where $\sigma_{\ln L_k|IM}$ is well above 1.0 (Figure 8-4a and Figure 8-4b). Using the lognormal assumption to compute $\sigma_{L_k|IM}$ results in an accurate prediction at large *IM* levels (i.e. as $\sigma_{L_k|IM}$ approaches a constant value), however, for smaller *IM* values it significantly over-predicts $\sigma_{L_k|IM}$ compared with the direct results. This over-prediction occurs because of the over prediction of $\sigma_{\ln L_k|IM}$ as shown in Figure 8-4a and Figure 8-4b, which is then amplified through the $e^{\sigma^2_{\ln L_k|IM}}$ term in Equation (8-19).

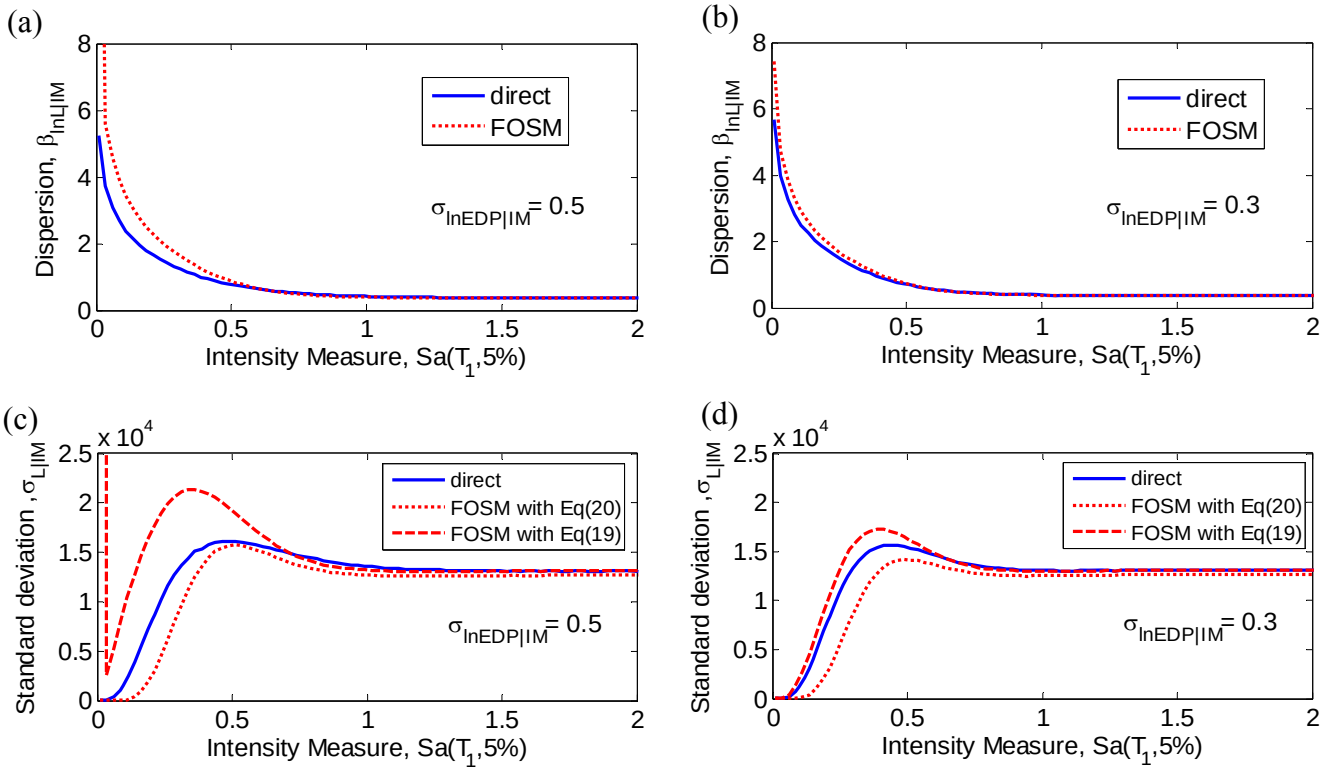


Figure 8-4: Accuracy of the FOSM method for: (a)&(b) dispersion; and (c)&(d) standard deviation in the prediction of the uncertainty in the loss given intensity relationship for the column component at two different levels of *EDP*|*IM* uncertainty.

In the case of Figure 8-4c, the dispersion using the FOSM method at small *IM* values is so large that using Equation (8-19) gives a 'non-convergent' value for $\sigma_{L_k|IM}$ (i.e. the value of $\sigma_{L_k|IM}$ does not converge to zero as $\mu_{L_k|IM}$ tends to zero). This does not occur in Figure 8-4d since the value of $\sigma_{\ln EDP|IM}$ is smaller than for Figure 8-4c, meaning the error in the approximation of the dispersion is not as significant. For this reason of 'non-convergence' in $\sigma_{L_k|IM}$ using the lognormal assumption (Equation (8-19)), the first-order assumption (Equation (8-20)) will be used in the remainder of this chapter.

Figure 8-5 illustrates the FOSM approximations for the standard deviation in the *L*|*IM* relationship for the partition component due to the onset of DS_2 (i.e. analogous to the results in Figure 8-3 for the expectation). Similar to the expected *L*|*IM* curves investigated previously it is evident that the approximation is worse for the partition component than the column component due to the small uncertainty in the partition fragility curves. Unlike the column component however, it is observed that around $IM = 0.15$ the standard deviation predicted using the FOSM method is quite significantly larger than that obtained via direct numerical integration. It is to be noted that $IM = 0.15$ is where the mean loss, $\mu_{L|IM}$, is 50% of

its absolute value (e.g. see Figure 8-3c and Figure 8-3d). Careful inspection of the standard deviation in loss given intensity for the column component (Figure 8-4c and Figure 8-4d) illustrates that there is also 'localised peak' around $IM = 0.5$, which is where the mean loss is 50% of its absolute value (see Figure 8-3a and Figure 8-3b). Therefore this localised peak is evidently occurring about the *IM* which causes the median *EDP* of the damage state in the component.

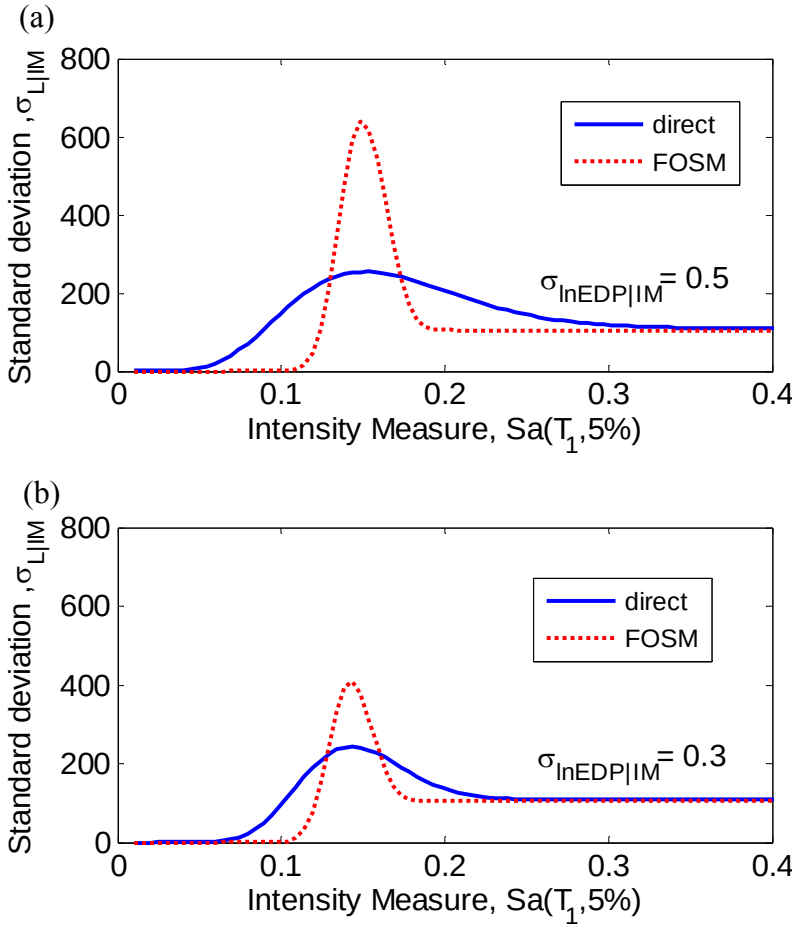


Figure 8-5: Accuracy of the FOSM prediction for the standard deviation in the loss given intensity relationship for the partition component for: (a) $\sigma_{\ln EDP|IM} = 0.5$; and (b) $\sigma_{\ln EDP|IM} = 0.3$.

8.5.3 Covariance in loss given intensity measure, $\sigma[L_1,L_2|IM]$

Equation (8-5) illustrates that in order to compute the standard deviation in the total loss, it is necessary to compute the covariance in the losses between different components for a given level of *IM*. As illustrated by Equation (8-7) this covariance computation requires knowledge of correlations within the *L*|*DS*, *DS*|*EDP*, and *EDP*|*IM* relationships. Such correlations occur for a variety of reasons, such as the similar repair actions required between similar components (affects *L*|*DS* correlation); components made of similar materials (affects

DS|*EDP* correlation), and components located in close proximity within the structure (affects *EDP*|*IM* correlation). As Equation (8-7) involves double integration it will be shown later that estimation of the covariance is computationally intensive. A FOSM approximation for the covariance of the loss between two components for a given intensity can be shown to be equal to [5]:

$$\sigma^2_{\ln L_k|IM,\ln L_{k'}|IM} \cong \sigma^2_{\ln EDP_i|IM,\ln EDP_{i'}|IM}\left(\frac{\partial \ln L_k|\ln EDP_i}{\partial \ln EDP_i}\right)^2\left(\frac{\partial \ln L_{k'}|\ln EDP_{i'}}{\partial \ln EDP_{i'}}\right)^2\Bigg|_{\mu_{\ln EDP_i|IM},\mu_{\ln EDP_{i'}|IM}} + \sigma^2_{\ln L_k|\ln EDP_i,\ln L_{k'}|\ln EDP_{i'}}\Big|_{\mu_{\ln EDP_i|IM},\mu_{\ln EDP_{i'}|IM}} \quad (8\text{-}21)$$

where all terms (derivatives and variances) are computed at $\mu_{\ln EDP_i|IM}$ and $\mu_{\ln EDP_{i'}|IM}$. The variance given by Equation (8-21) can then be converted to its non-log form by:

$$\sigma^2_{L_k|IM,L_{k'}|IM} \cong \mu_{\ln L_k|IM}\mu_{\ln L_{k'}|IM}\sigma^2_{\ln L_k|IM,\ln L_{k'}|IM} \quad (8\text{-}22)$$

Due to space limitations, a direct comparison of the FOSM and direct computations of the covariance between two components is not given here. The covariance of the loss between two components is however directly related to the standard deviation in the loss of each of the components and their correlation. As generally the FOSM method leads to an under prediction of the standard deviation in loss given intensity (the exception being for components with small uncertainty in the DS fragility, in which an over approximation occurs about the *IM* causing the median *EDP* of the fragility function) it can be inductively stated that the covariance will also likely be under predicted (as positive correlations occur in general).

One further point regarding the computation of the covariance in the loss between two components is also warranted. As previously mentioned, the computation of the correlation between losses in different components is a difficult task with a vast lack of data. As such, one may wish to consider upper and lower bounds, by considering no correlations and perfect correlations between components. In the case of no correlations, all covariance terms will vanish and the variance of the total loss will simply be the sum of the variances of each of the components. In the case of assuming perfect correlations between *L*|*DS*, *DS*|*EDP*, and *EDP*|*IM* relationships for two components that are identical (i.e. means and variances for each of the relationships are equal), the covariance computation (Equations (8-6) and (8-7)) will reduce to the product of the standard deviations in the loss for each of the components (Equation (8-3)). This result is, in general, not true for two different components which have

perfect correlations within the *L|DS*, *DS|EDP*, and *EDP|IM* relationships. That is, in general, even if perfect correlations exist (for two different components) between the three deaggregated (*L|DS*, *DS|EDP*, and *EDP|IM*) relations the correlation between the *L|IM* relations will not necessarily be one. This point is illustrated for the case study considered later in the chapter.

8.5.4 Epistemic variance in collapse fragility curve, $\sigma^2[P(C|IM)]$

The FOSM method can also be used when computing the epistemic uncertainties in a cumulative distribution, due to uncertainty in its mean value. In loss estimation methodologies, such a case arises when computing the uncertainty in: (i) the cumulative probability of collapse given *IM* (for computing the collapse hazard); (ii) the probability of exceeding a given level of demand for a given *IM* (for computing the demand hazard); and (iii) the probability of exceeding a given level of loss for a given *IM* (for computing the loss hazard). The demand and loss hazard cases are further complicated in that the cumulative probability distribution needs to also account for the probability of collapse. Details on the methodological formulation to include collapse can be found in Baker and Cornell [4]. As the cumulative distribution (which we will assume to be lognormal [11]) for the collapse probability has less of the aforementioned complications we will use it here to illustrate the accuracy of the FOSM method.

Using Bayes' Theorem [8], it is possible to compute the mean and variance in the probability of collapse given *IM* due to uncertainty in the mean *IM* causing collapse from:

$$\mu_{P(C|IM)} = \int P\left(C \mid IM, \mu_{IM|C}\right) f\left(\mu_{IM|C}\right) d\mu_{IM|C} \tag{8-23}$$

$$\sigma^2_{P(C|IM)} = \int \left[P\left(C \mid IM, \mu_{IM|C}\right)\right]^2 f\left(\mu_{IM|C}\right) d\mu_{IM|C} - \mu^2_{P(C|IM)} \tag{8-24}$$

where *C* denotes 'collapse'; and $\mu_{IM|C}$ is the mean *IM* causing collapse which is an uncertain quantity. For the particular case of the probability of collapse being a lognormal random variable then FOSM approximation for Equation (8-24) is [5]:

$$\sigma^2_{P(C|IM)} \cong \phi\left(\frac{\ln(im) - \overline{\mu}_{\ln IM|C}}{\sigma_R}\right)^2 \frac{\sigma^2_{\mu_{\ln IM|C}}}{\sigma^2_R} \tag{8-25}$$

where σ_R is the aleatory uncertainty in the collapse fragility curve; $\sigma_{\mu_{\ln IM|C}}$ is the epistemic uncertainty in the (logarithmic) mean *IM* causing collapse, $\mu_{\ln IM|C}$; and $\overline{\mu}_{\ln IM|C}$ denotes the

mean (with respect to epistemic uncertainty) estimate of $\mu_{\ln IM|C}$.

Figure 8-6a illustrates the accuracy of the FOSM method for relatively typical values of the aleatory and epistemic uncertainties (actual values are usually in the range 0.4-0.5 [12]). It is observed that for *IM* values near the mean *IM* causing collapse the FOSM approximation is too large relative to the direct computation, while for *IM* values at several standard deviations from the mean the FOSM approximation actually falls below the direct values. Figure 8-6b illustrates a slightly more unusual scenario with large aleatory and small epistemic uncertainties. Despite the small value of the epistemic uncertainty in this case, the prediction of the FOSM approximation of the direct solution obtained via numerical integration is poor (note however that the absolute value of the uncertainty in Figure 8-6b is small relative to Figure 8-6a).

8.6 Application to loss assessment of a structure

In this section the FOSM method is applied in the loss assessment of a 10 storey reinforced concrete moment frame structure, and compared to the results obtained via direct integration of the loss assessment governing equations. A detailed discussion on the case study structure and its loss assessment is given in [7], and briefly explained below.

8.6.1 Case study structure

The case study structure used herein to illustrate the accuracy of the FOSM method for the loss estimation of an entire structure is based on the Red Book building [13] which acts as a design example of the New Zealand Concrete Code [14]. The primary lateral load carrying system consists of four one-way perimeter moment resisting frames which are 3 bays long. Vertical loads are transferred primarily through interior columns with gravity beams supporting one-way floor units. The loss assessment considers 115 different component groups which encompass structural components, non-structural components and contents. Only losses due to direct repair cost of damage were considered (i.e. human casualties and business disruption were not considered), and epistemic uncertainties were not considered.

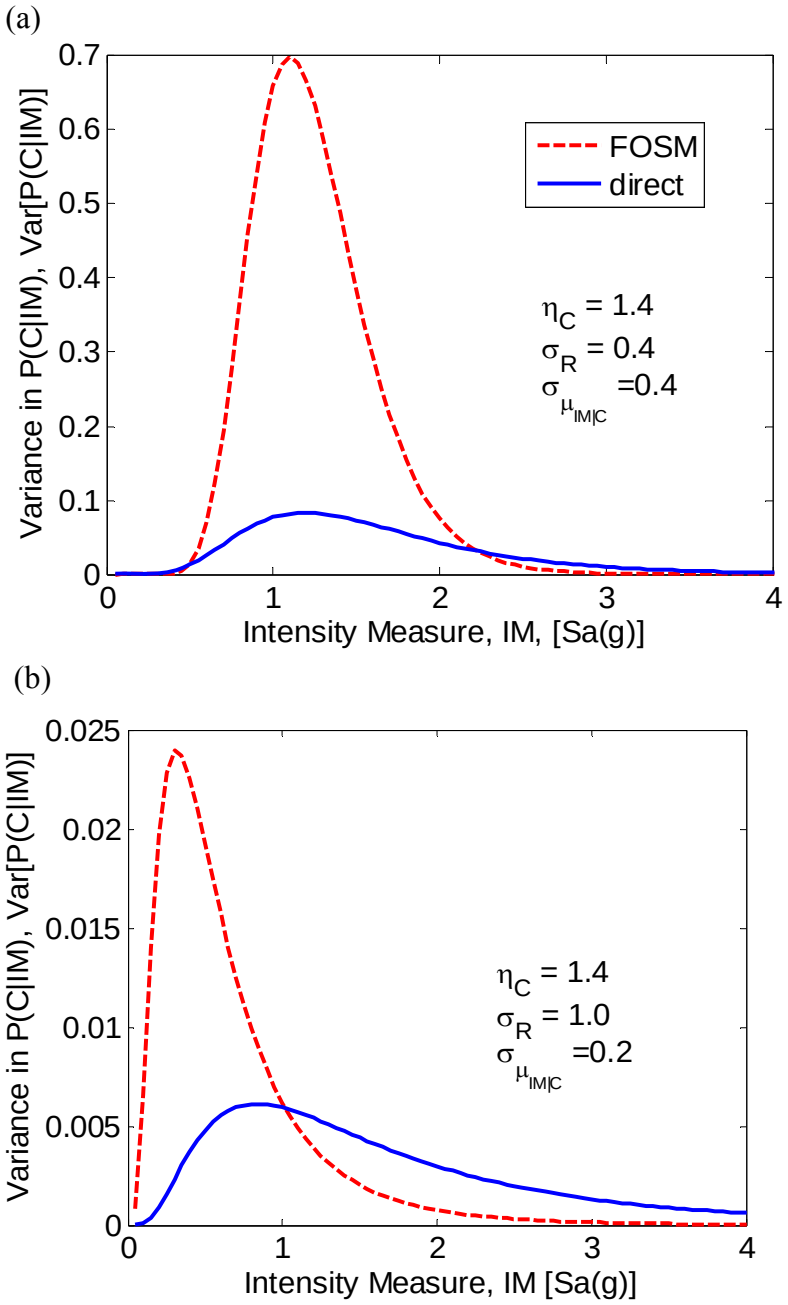


Figure 8-6: Error in the prediction of the epistemic variance in the cumulative distribution for collapse probability using the FOSM method using: (a) $\eta_C = 1.4$; σ_R=0.4; $\sigma_{\mu IM|C} = 0.4$; and (b) $\eta_C = 1.4$; σ_R=1.0; $\sigma_{\mu IM|C} = 0.2$. Here η_C is the median *IM* value causing structural collapse.

8.6.2 Approximate methods of uncertainty propagation used in loss assessment

As has been previously discussed, there are several different options for the inclusion of the FOSM method in the seismic loss assessment computations. Here we will compare two different alternatives, which have been used by previous researchers (but not compared

against the exact solution).

Baker and Cornell [5] propose using the FOSM method to obtain the mean and standard deviation of loss given *IM* for individual components, and for computing the covariance in the *L*|*IM* relationship between different components when computing the standard deviation for the total loss given *IM*. The distribution of the total loss is then combined with the ground motion hazard via numerical integration. Aslani [6] uses the FOSM method only for the computation of the covariance terms when determining the standard deviation in the total loss. That is, Aslani [6] computes the mean and standard deviation of the loss given intensity for a single component using direct numerical integration. For brevity, reference to the approach proposed by Baker and Cornell [5] is denoted as the 'FOSM' approach, while the approach of Aslani [6] is denoted as the 'partial' approach. Solution of the problem via direct numerical integration is denoted the 'direct' approach.

The seismic loss assessment tool, SLAT [15] is used for conducting the loss assessment. In particular, the computational scheme for directly performing the numerical integrations utilizes the magnitude-oriented adaptive quadrature (MAQ) algorithm [16]. The integration parameters used were an error tolerance of 0.005 (0.5%) and a maximum number of function evaluations of 300 (for all computations presented here integral convergence was achieved).

8.7 Expected loss given *IM*

Earlier in the chapter, the error associated with the FOSM method for computation of the expected loss given intensity for a single component was illustrated. Three important observations regarding the error in the FOSM approximation were that: (i) the error increased as the dispersion in the *EDP*|*IM* relationship increased; (ii) the error increased as the dispersion in the fragility functions decreased; and (iii) the FOSM method under and over predicted the loss for intensities below and above the intensity causing a 50% probability of exceedance of the damage state, respectively.

Figure 8-7a illustrates the computation of the total expected loss given intensity and no collapse ($\mu_{L_T|IM,NC}$) for the entire structure. It is first noted that the partial method yields the same results as the direct method since the expected total loss given no collapse is simply the sum of the expected loss given intensity for each component (Equation (8-6)). Secondly, it is noted that while the FOSM method under predicts the expected loss relative to the direct solution its approximation is surprisingly good relative to the results observed for the

approximation of the expected loss given intensity for single components (i.e. Figure 8-3). The reason for this improved performance for the total loss (compared to the loss for a single component) can be attributed to the third point noted above whereby the negative errors in the estimation of the loss in one component are offset by positive errors in the estimation of loss for another component. Figure 8-7b illustrates the relative error between the FOSM and direct solutions for the total expected loss as a function of *IM*. The two different lines illustrate the error in the expected total loss given no collapse, $\mu_{L_T|IM,NC}$, and the expected total loss after conditioning on collapse is removed, $\mu_{L_T|IM}$ [7]. As both the FOSM and direct approaches consider the loss due to global collapse in the same fashion (by combing the collapse fragility curve, $P_{C|IM}$, with the mean loss due to collapse, $\mu_{L_T|C}$) then as *IM* increases, and a significant portion of the expected loss is due to collapse (due to an increasing probability of collapse), the relative error will tend to zero as $P_{C|IM}$ approaches unity.

Figure 8-7b illustrates that for $IM > 0.3$g the FOSM approximation results in less than a 10% under prediction of the expected loss, and the error is below 5% for $IM > 0.5$g (this corresponds to a ground motion with approximately a 2% exceedance in 50 years). The fact that the error is quite large for low levels of ground motion is that most components in the structure are subjected to demands which give less than a 50% probability of damage state exceedance. As it was previously observed that the FOSM method under predicts the loss in individual components for low *IM* levels, then for these low *IM* levels, the FOSM method under predicts the expected loss for the majority of components, and therefore the amount of subtractive cancellation is small. When the ground motion hazard curve and the expected loss as a function of intensity, $\mu_{L_T|IM}$, are convoluted (using direct numerical integration) to obtain the expected annual loss, EAL (i.e. Equation (8-9)), it was found that the FOSM method gives an under-approximation of the EAL by 15% compared to that obtained via direct numerical integration.

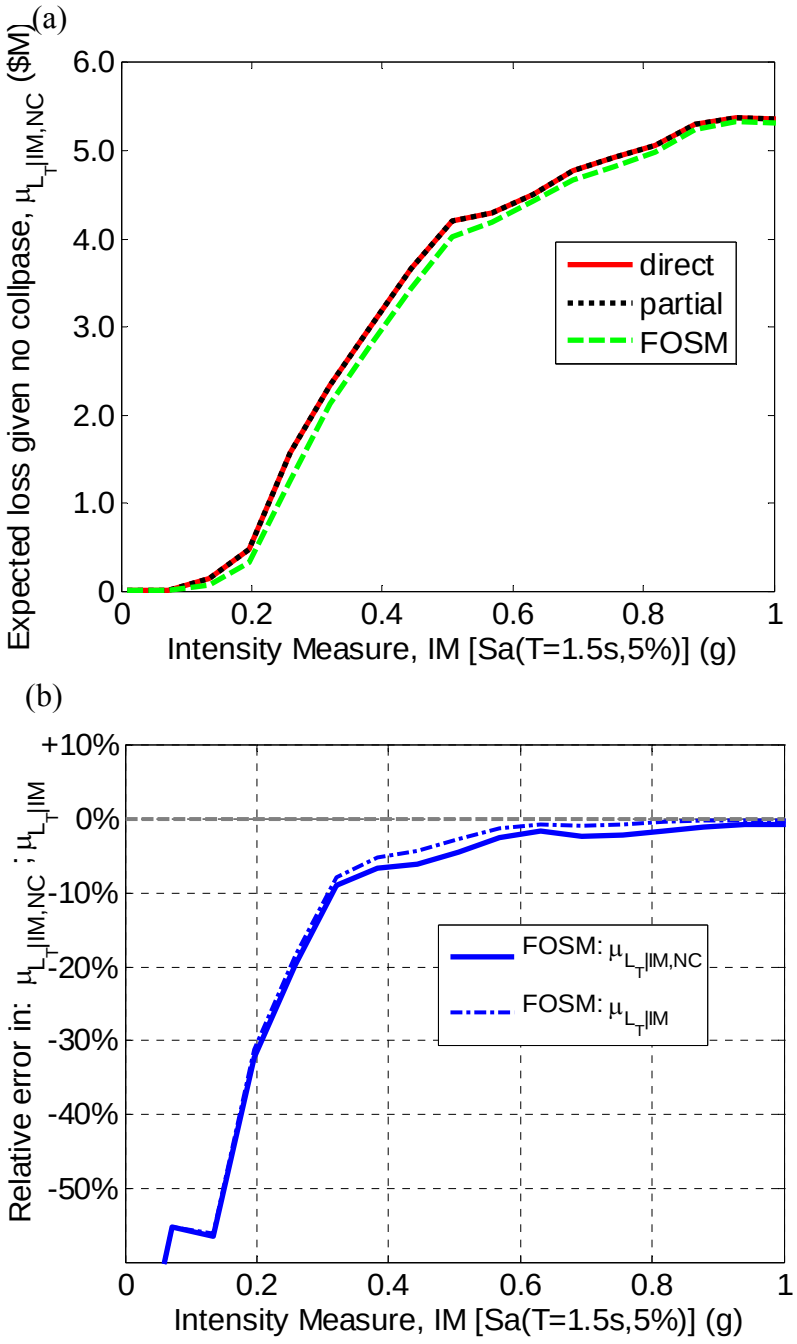


Figure 8-7: (a) Expected total loss given intensity for the case study structure; and (b) relative error of the FOSM approximation as a function of *IM*.

For this specific structure, over the majority of *IM* levels, the loss is not dominated by a single type of component. Therefore the aforementioned subtractive cancellation effect means that the FOSM method is relatively accurate for this example. Care should therefore be taken on a case-by-case basis that the loss (for a given level of intensity) is not dominated by a single component (or component type). If the loss is dominated by a single component then the accuracy of the FOSM method will likely be reduced, similar to that observed for the parametric study presented earlier for individual components (see Figure 8-3).

Caution should also be exercised when investigating the deaggregation of loss given *IM* for use in decision making purposes, because as mentioned above, the contribution of various types of components may be significantly different when using the FOSM method in place of direct numerical integration. Figure 8-8 illustrates the deaggregation of the expected loss given no collapse for *IM* = 0.2g *Sa* (which has a rate of exceedance of 17% in 50 years for the site), by fragility type computed using both the direct and FOSM approaches. It is apparent in Figure 8-8 that while the hierarchy of component contribution to the total loss does not change significantly, there are a few components whose contribution is markedly different between the two solutions. Changes in contributions to the loss of particular note are: server and network equipment reduce from 18% to 13%; acoustical ceiling reduce from 9% to 5%; paint increases from 9% to 15%. These increasing and decreasing proportions can be explained by recalling from Figure 8-4 that the FOSM method under approximates the loss for IM values below that causing the median loss, and over approximates the loss for IM values above the median loss.

8.7.1 Standard deviation in loss given *IM*

Figure 8-9 illustrates the comparison in the computation of the standard deviation in the total loss. As it is beyond the scope of this work to delve deeply into the consideration of correlations in the loss assessment, we will merely consider the upper and lower bounds of perfect and no correlations, respectively. Figure 8-9a shows the accuracy of the FOSM method for predicting the standard deviation for the no correlation case. Note that since the partial method computes the mean and standard deviation in the loss for individual components using numerical integration, then for the case of no correlations (i.e. the covariance terms in Equation (8-5) are zero) the standard deviation in the total loss is estimated exactly. Similar to the results observed in the prediction of the standard deviation in the loss for individual components, the standard deviation in the total loss is under predicted by the FOSM method over the entire range of *IM*. Figure 8-9c illustrates the relative error in the FOSM method for the case of no correlations. In particular, while there is quite a significant difference in the prediction of the standard deviation in the total loss conditioned on no collapse, once conditioning on no collapse is removed (i.e. Equation (8-8)), the error reduces significantly. This is due to the fact that for the no correlation case the standard deviation in the total loss given no collapse is small relative to the standard deviation in the total loss given collapse (which is computed exactly).

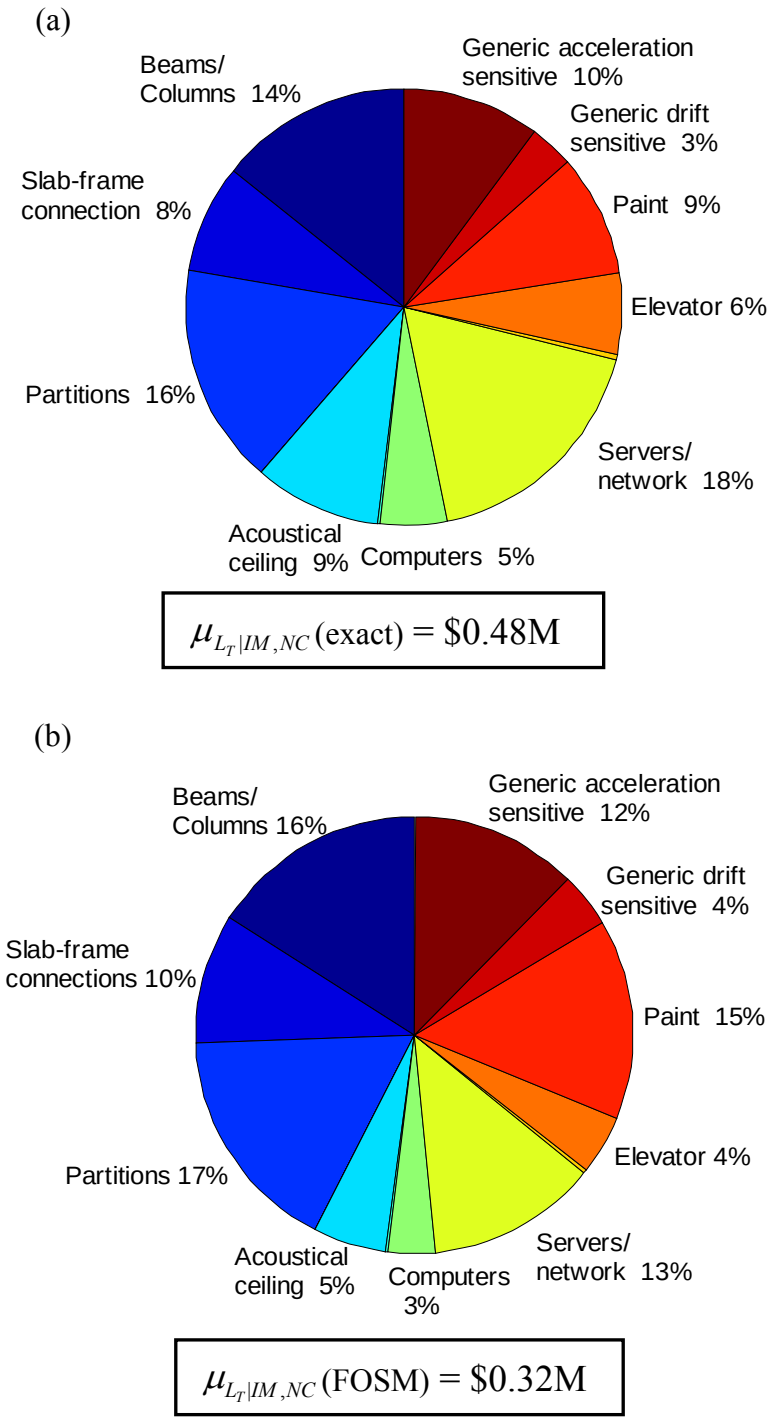


Figure 8-8: Deaggregation of the expected total loss given no collapse at IM = 0.2g Sa,by fragility type using: (a) direct numerical integration; and (b) FOSM method. *Components with less than 1% contribution have not been annotated

Figure 8-9b illustrates the prediction of the standard deviation in the total loss given no collapse for the case of perfect correlations in the $L|DS$, $DS|EDP$, and $EDP|IM$ relationships. Note that as mentioned previously, perfect correlations in each of the aforementioned three relations does not imply a perfect correlation within the $L|IM$ relationship. The error in making this assumption is illustrated in Figure 8-9b, with a relative over-prediction of

approximately 20% (Figure 8-9d) in the case of no collapse. Figure 8-9b also illustrates the accuracy of the FOSM and partial methods in predicting the standard deviation in the total loss given no collapse. It should be noted that the difference between the two methods is the computation of the variance terms (first part of Equation (8-5), which the partial method solves exactly, while the FOSM method uses Equation (8-18)). The computation of the covariance terms (second part of Equation (8-5)) is approximated using Equation (8-21) in both FOSM and partial methods. From Figure 8-9b it is apparent that both methods quite significantly under predict the magnitude of the standard deviation.

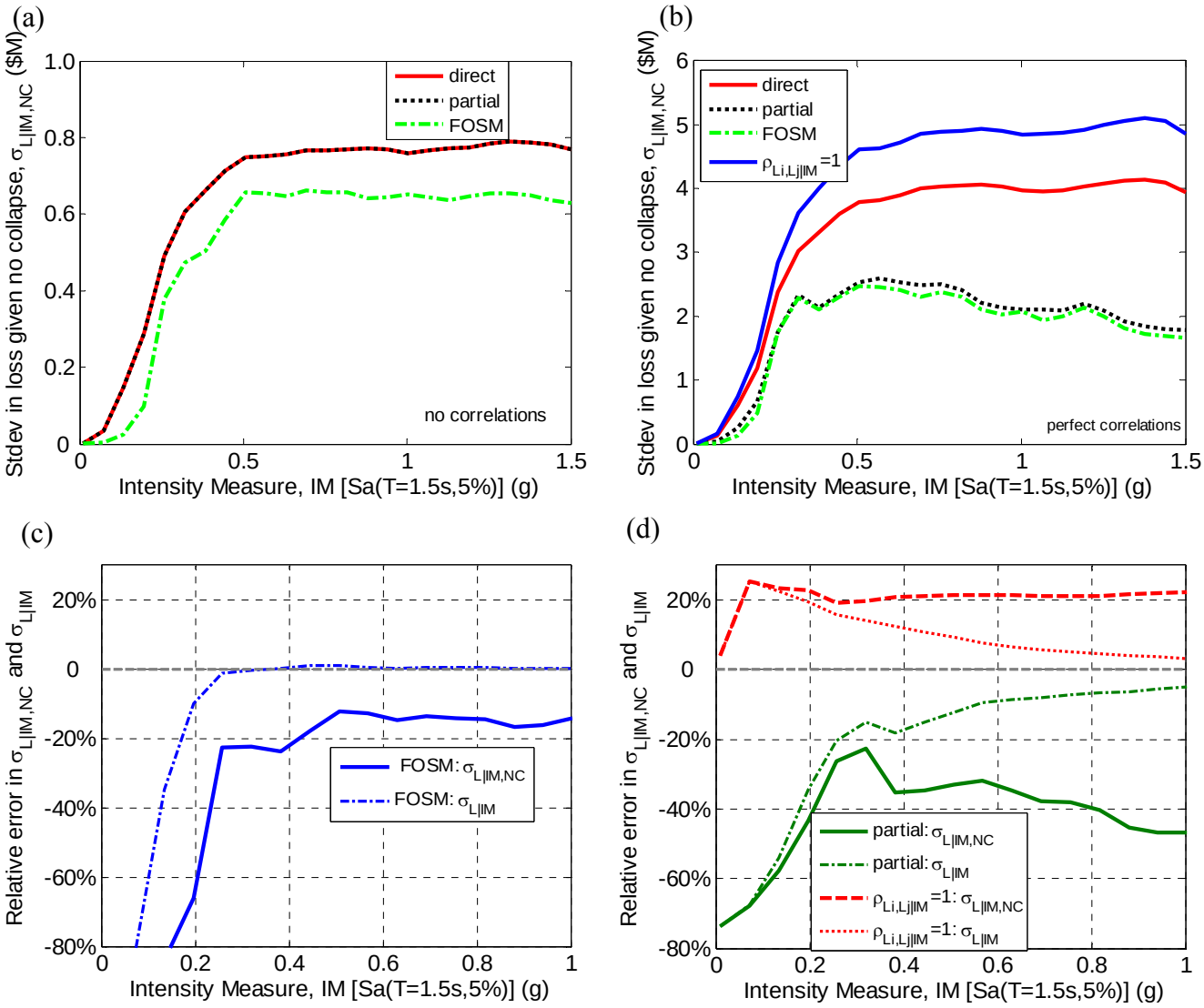


Figure 8-9: Comparison of the computation of the standard deviation in the total loss: (a) assuming no correlations; (b) assuming perfect correlations; (c) relative error of the FOSM method when assuming no correlations; and (d) relative error of FOSM and partial methods when assuming perfect correlations.

Figure 8-9d illustrates that the relative error in the standard deviation using the partial method is always greater than 20% for the case of no collapse (the error in the FOSM method

is not shown, because of its similarity to the partial method). While this error is reduced when collapse is considered, the reduction is not as marked as that of the case of no correlations, with the error greater than 20% for $IM < 0.25$g and less than 10% for $IM > 0.5$g. This is because the magnitude of the standard deviation given no collapse is about 5 times larger than the case of no correlations (and so the collapse standard deviation does not dominate). Figure 8-9d also illustrates the relative error in assuming that if perfect correlations exist between the *L*|*DS*, *DS*|*EDP*, and *EDP*|*IM* relationships, then a perfect correlation exists for the *L*|*IM* relationship. This assumption causes an over-approximation of approximately 20% in the standard deviation of the loss given no collapse.

8.7.2 Loss hazard

The loss hazard curve, which gives the annual rate of exceeding various levels of loss, can be obtained from the distribution of the total loss and the ground motion hazard curve, as given by Equation (8-10). Figure 8-10a and Figure 8-10b illustrate the errors in the computation of the loss hazard curve for the two cases of zero and perfect correlations, respectively. For the case of no correlations (Figure 8-10a), since the error in the standard deviation using the FOSM method (considering both collapse and non collapse cases) is below 10% once $IM > 0.3$g then the reason for the difference between the direct and FOSM-based loss hazard curves is primarily due to the difference in the computed expected loss given *IM* (Figure 8-7). For example, at an exceedance rate of $2\text{x}10^{-3}$, values of \$1.3 M, and \$1.6 M are obtained for the FOSM and direct methods, respectively. Again it is noted that in the case of no correlations the partial approach is exact and therefore is not shown here.

Figure 8-10b illustrates the error in the loss hazard curve in the case of perfect correlations. It can be seen that the under prediction of the standard deviation in the total loss by the partial method (and FOSM method) results in an un-conservative loss hazard curve. The conservative assumption of perfect *L*|*IM* correlations is also shown, which results in a conservative loss hazard curve. For an exceedance rate of $2\text{x}10^{-3}$, values of \$0.87 M, \$1.08 M, and \$1.06 M are obtained for the partial, direct, and perfect *L*|*IM* correlation cases, respectively. Comparison of Figure 8-10a and Figure 8-10b also illustrates the significant effect of uncertainty in the *L*|IM relationship on the shape of the loss hazard curve, and the rate of occurrence of various levels of loss, particularly at low rates of exceedance.

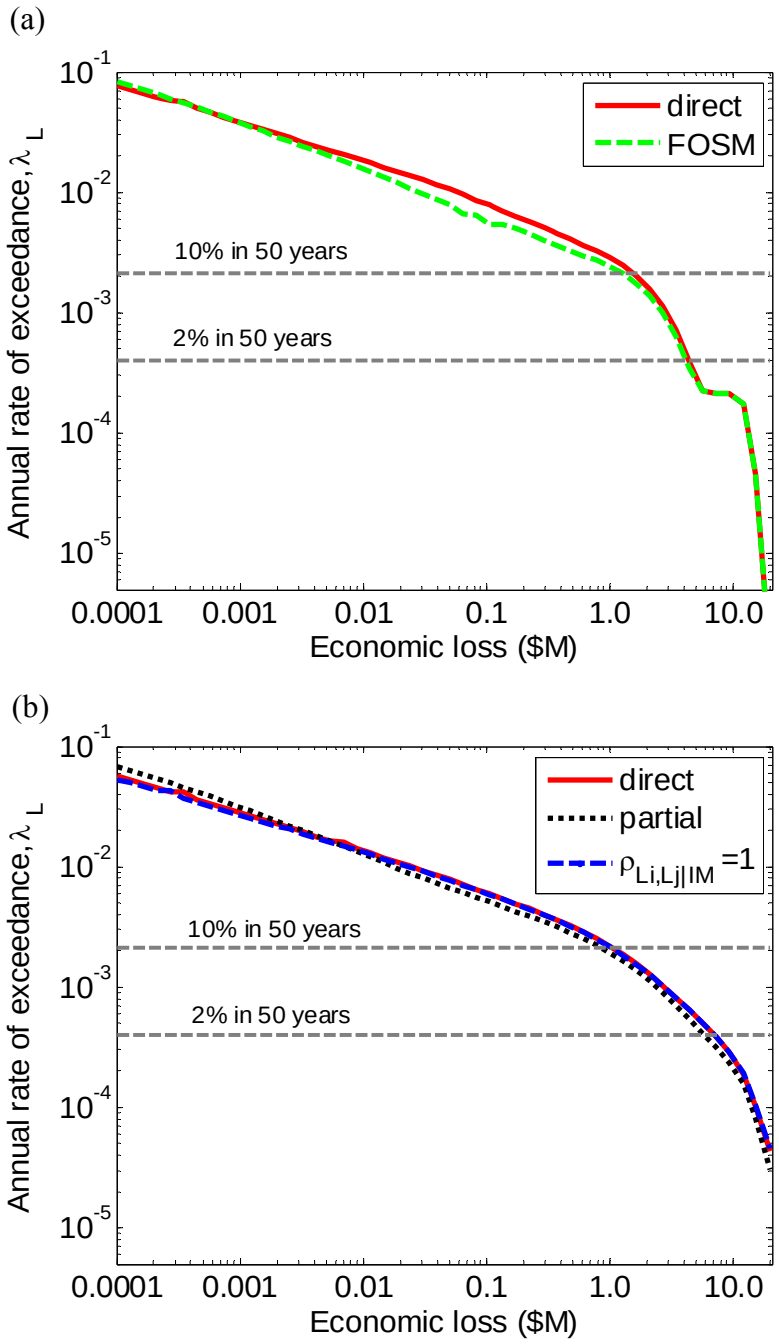


Figure 8-10: comparison of the computation of the loss hazard curve: (a) assuming no correlations; and (b) assuming perfect correlations.

8.7.3 Computational demand

The primary benefit with any approximate method is its reduction in effort to the user, whether it be computational as in this case, or expertise required in the input information. Thus, in order to put the accuracy of the aforementioned approximate methods into perspective, it is necessary to focus on their benefits of computational reduction. The case study presented in the previous section represents a good example to illustrate the

computational demands in performing a loss assessment, with a total of 115 different components monitored over 21 different *EDP*'s.

Table 8-2 presents the computational times required when in performing the seismic loss assessment on a Pentium 4 processor with 3.0 GHz CPU and 512 MB RAM using the seismic loss assessment tool (SLAT) [15] with 25 output points (i.e. 25 *IM* points for the *L*|*IM* relations, or 25 Loss points for the loss hazard). The computations use the MAQ integration algorithm [16] with an error tolerance of 0.001 (0.1%). For comparison the computational demand when computing the loss hazard curve with a larger (1%) error tolerance is also given. Table 8-2 illustrates primarily two key points, the first being that computation of the loss hazard curve is significantly more demanding than that for the *L*|*IM* and *EAL* computations. Secondly, the effect of considering correlations is also significant in increasing the computational demand. The reason for the latter observation is explained in the following sentences. If we consider the computational effort of the mean and standard deviation of loss given *IM* for a single component to be $2\,\overline{n}_{feval}$, where $\overline{n}_{feval}$ is the average number of function evaluations required to perform the integration (in Equation (8-2)), and the coefficient '2' is for computation of both the mean and standard deviation. Then the total computational work to obtain the mean and standard deviation of the total loss for no correlations is approximately $2\,\overline{n}_{feval}\,m$, where m is the number of components. When correlations are considered the covariance in the loss given *IM* between two different components (Equations (8-6) and (8-7)) requires approximately $\overline{n}^2_{feval}$ function evaluations, and the second term in Equation (8-5) indicates that for m components there are $\frac{1}{2}m(m-1)$ covariance computations to be computed. Thus the ratio between the computational effort to evaluate the covariance terms in Equation (8-5) and the remainder of the terms in Equations (8-4) and (8-5) is given by:

$$\frac{Comp\left[\sum\sum \sigma_{L_i|IM,L_j|IM}\right]}{Comp\left[\sum \mu_{L_i|IM}, \sum \sigma^2_{L_i|IM}\right]} \sim \frac{\frac{1}{2}m(m-1)\overline{n}^2_{feval}}{2\overline{n}_{feval}m} = \frac{\overline{n}_{feval}(m-1)}{4} \tag{8-26}$$

Equation (8-26) illustrates that for the case study structure with m = 115 components, and given that $\overline{n}_{feval}$ ~60 the computational effort to compute the covariance terms is of the order of 1700-times that to compute the other terms in the *L*|*IM* relation. For the special case of perfect correlations within the *EDP*|*IM* relationship, the computational demand ratio reduces to approximately $(m-1)/4$. As Equation (8-26) illustrates that the computational work to determine the covariance terms is a quadratic relationship of the number of

components, *m*, then it is noted that the use of deaggregation to determine if any components offer insignificant contribution to the loss estimation can significantly reduce the computational demand. For the case study presented in the previous section, up to 15 components could have been removed with a less that 2% effect on the results, which would reduce the computational demand by approximately 38%.

Table 8-2: Computational times (in seconds) for performing seismic loss estimation.

	No correlations			Perfect correlations		
Relationship	FOSM	Partial	Direct	FOSM	Partial	direct
$\mu_{L_T\|IM}$, $\sigma^2_{L_T\|IM}$ (Eq. 6-10)	0.64	1.67	1.67	4.8	5.8	78.2
μ_{L_T} (Eq 11)	2.47	15.0 s	15.0	43.8	53.1	705
λ_{L_T} (Eq 12)	11.2 (3.97)*	103 (18.3)*	103 (18.3)*	391 (119)*	480 (142)*	6200 (1205)*

*Computation times using an error tolerance of 1% in brackets.

8.8 Conclusions

The efficacy of approximate methods of uncertainty propagation in seismic loss estimation has been investigated in this chapter. The approximate methods are based on the use of the first-order second-moment (FOSM) method which uses a first-order Taylor Series approximation to compute the first two moments of functions of random variables. The FOSM method was used primarily to determine the moments of the loss given intensity measure (*L*|*IM*) relationship. It was illustrated that the accuracy of the FOSM method is directly related to the uncertainty in the *EDP*|*IM* relationship, with increasing error as the uncertainty increases. As the error in the approximation of the mean loss given intensity for a given component can be either an over or under approximation, then when computing the mean loss given intensity for the entire structure, some cancellation errors occur, with relative errors up to 50% for small *IM* values and reducing error with increasing *IM* as more components contribute to the total loss. The same cannot be said for the approximation of the standard deviation of the loss given intensity, which was almost always under predicted by more than 20% for the case-study structure considered. Despite the cancellation errors giving a small total error in the mean loss given intensity for the entire structure, care should be taken in deaggregation of the results, which tend to have significantly larger errors compared to the exact solution. The effects of the errors in the *L*|*IM* relationship on the resulting loss hazard

curve were also investigated. Computational times to conduct the analysis on a standard PC indicate that the loss hazard computation was approximately 8- and 80-times more computationally demanding than computation of the *L|IM* and expected annual loss (*EAL*) decision variables, and that consideration of component correlations is also demanding. Both approximate and exact methods have their pros and cons in various situations. This chapter therefore elucidates the situations in which an analyst may lean toward one of the methods over the other, be it based on computational demands or accuracy tolerance.

8.9 Acknowledgements

Financial support of the first author by the New Zealand Tertiary Education Commission is greatly appreciated. Discussions between Dr. Jack Baker and the first author are gratefully acknowledged.

8.10 References

[1] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).

[2] Deierlein GG, Krawinkler H, and Cornell CA. A framework for performance-based earthquake engineering, in *7th Pacific Conference on Earthquake Engineering*, Christchurch, New Zealand, 2003.

[3] Cornell CA. A probability based structural code. *Journal of the American Concrete Institute* 1969; **66**(12): 974-985.

[4] Baker JW and Cornell CA. Uncertainty Specification and Propagation for Loss Estimation using FOSM Method. Peer Report No 2003/07, 2003. 100. http://peer.berkeley.edu/publications/peer_reports.html

[5] Baker JW and Cornell CA. Uncertainty propagation in probabilistic seismic loss estimation. *Structural Safety* 2008; **30**(3): 236-252.

[6] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[7] Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, and Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society for Earthquake Engineering* 2009; **42**(2): 96-110.

[8] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[9] Melchers RE. *Structural Reliability Analysis and Prediction*. John Wiley and Sons: Chichester, 1999;

[10] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[11] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1901-1914.

[12] Haselton CB. Assessing Collapse Safety of Modern Reinforced Concrete Moment Frame Buidlings. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2007, 312.

[13] Bull DK and Brunsdon D. Examples of Concrete Structural Design to New Zealand Standards 3101. New Zealand, 1998.

[14] Standards New Zealand. NZS 3101 1995: Part 1: Concrete Structures Standard. Wellington, NZ, 1995.

[15] Bradley BA. User manual for SLAT: Seismic Loss Assessment Tool version 1.14. *University of Canterbury Research Report No.2009-01*,Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand, 2009. 94.

[16] Bradley BA, Lee DS, Broughton R, and Price C. Efficient Evaluation of Performance-based Earthquake Engineering Equations. *Structural Safety* 2009; **31**(1): 65-74.

9. Component Correlations in Structure-Specific Seismic Loss Estimation

Bradley BA, Lee DS. Component correlations in structure-specific seismic loss estimation. *Earthquake Engineering and Structural Dynamics* 2009 (in press)

9.1 Abstract

This chapter addresses correlations between multiple components in structure-specific seismic loss estimation. To date, the consideration of such correlations has been limited by methodological tractability; increased computational demand; and a paucity of data for their computation. The effect of component correlations, which arise in various forms, is however a significant factor affecting the results of structure-specific seismic loss estimation and therefore it is prudent that adequate consideration is given to their effect. This chapter provides details of a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered. Methods to determine the necessary correlations are discussed, particularly those that can be used in the absence of sufficient empirical data, for which values are suggested based on judgement. The effects of various assumptions regarding correlations are illustrated via application to a case-study office structure. It is observed that certain correlation assumptions can lead to errors in excess of 50% in the lognormal standard deviation in the loss given intensity and loss hazard relationships, while full consideration of partial correlations is 50-times more computationally expensive than other assumptions.

9.2 Introduction

Structure-specific seismic loss estimation, in line with the performance-based earthquake engineering (PBEE) framework, involves detailed consideration and computation of losses in a structure due to seismic risk. Uncertainties are explicitly accounted for by treating variables, such as seismic hazard and structural response, probabilistically and integrating over their range of possible values when computing decision variables useful in decision making. Accurate seismic loss estimation requires not only consideration of the uncertainties in the individual components which comprise the structure, but also the or correlations between the different components.

Seismic loss estimation methodologies which allow consideration of correlations and the effect of such correlations have received little attention in literature due to, in the author's opinion, several reasons: (i) inevitably their consideration significantly increases the complexity of the algorithms required to perform the loss estimation; (ii) the complicated algorithms significantly increase the computational demand to perform the loss estimation; (iii) there is likely a lack of appreciation for the influence of correlations in the results of the loss estimation.

In the seismic loss estimation framework discussed herein, there are three different correlations which exist between components; these are: (i) correlation between engineering demand parameters (EDP's) for a given intensity measure (IM); (ii) correlation between damage states (DS) given EDP*'s*; and (iii) correlation between component loss (L) given DS's. Recent efforts utilizing the Pacific Earthquake Engineering Research (PEER) centre framework, such as Goulet *et al*. [1] and Mitrani-Reiser [2] have considered only expected losses, and therefore neglected correlations (which, for their framework equations, only affects the variance in the loss). Aslani [3], Porter and Kiremidjian [4] , Iervolino *et al*. [5], and the ATC-58 guidelines being developed [6] consider correlations explicitly in the $EDP|IM$ relation from the results of seismic response analysis, but Porter and Kiremidjian [4] , Iervolino *et al*. [5], and ATC-58 [6] do not consider correlations in the $DS|EDP$ and $L|DS$ relationships. Aslani [3] and Lee and Kiremidjian [7] consider correlations in the discrete damage state variable using a cumbersome optimisation algorithm, while Baker [8] has recently commented on the use of a more flexible alternative approach. Aslani [3] also considered correlations in the $L|DS$ relationship using correlation data from the construction sector. Aslani [3] and Baker and Cornell [9] both use the first-order second-moment (FOSM)

approximation when computing the covariance terms in the total loss because of the perceived computational demand of direct numerical integration. The FOSM approximation has been shown to be of limited accuracy in computing such covariance terms compared to direct evaluation via numerical integration [10]. As far as the author is aware only Aslani [3] has briefly investigated the effect of correlation assumptions on the standard deviation in the total loss given intensity.

The intention of this chapter is four-fold. Firstly, the necessary details of a tractable and computationally efficient framework which can account for such correlations are presented. Secondly, significant discussion is given to the determination of the required correlation coefficients, both those which can be computed directly from empirical data, and methods which can be used in the absence of sufficient data. Thirdly, the interaction of correlations and epistemic uncertainties in loss estimation is discussed. Finally, the effects of various assumptions regarding the treatment of correlations are illustrated via an application to a typical office structure.

9.3 Seismic loss estimation methodology considering component correlations

9.3.1 General methodological details

This section presents the mathematical details of a seismic loss estimation method which explicitly accounts for correlations. The basis of the methodology is the PEER performance-based earthquake engineering (PBEE) framework [11] which has been employed by several other researchers [e.g. 1, 3, 9]. Below the various aspects of this framework which are affected by correlations are discussed.

The total loss incurred in a structure when subjected to a ground motion of a specified intensity measure (IM) is conditioned on the mutually exclusive and collectively exhaustive events of collapse and non-collapse. From the total expectation theorem [12], the mean, $\mu_{L_T|IM}$, and variance, $\sigma^2_{L_T|IM}$, of the total loss given *IM* are given by [e.g. 3]:

$$\mu_{L_T|IM}(im) = \mu_{L_T|IM,NC}(im)\left(1 - P_{C|IM}(im)\right) + \mu_{L_T|C} P_{C|IM}(im) \tag{9-1}$$

$$\sigma^2_{L_T|IM}(im) = \sigma^2_{L_T|IM,NC}(im)\left(1 - P_{C|IM}(im)\right) + \sigma^2_{L_T|C} P_{C|IM}(im) + \left(\mu_{L_T|IM,NC}(im) - \mu_{L_T|C}\right)^2 P_{C|IM}(im)\left(1 - P_{C|IM}(im)\right) \quad (9\text{-}2)$$

where $\mu_{L_T|IM,NC}(im)$, $\sigma^2_{L_T|IM,NC}(im)$, $\mu_{L_T|C}$, and $\sigma^2_{L_T|C}$ are the mean and variance in the loss given *IM* = *im*, no collapse and collapse, respectively; $P_{C|IM}(im)$ is the probability of collapse given *IM* = *im*. Given collapse occurs it is assumed that the mean and variance in the total loss are independent of *IM* (i.e. $\mu_{L_T|IM,C} = \mu_{L_T|C}$).

In the case of no collapse, the total loss is comprised of the sum of the loss to individual components at spatially different locations throughout the structure, with mean, $\mu_{L_T|IM,NC}$, and variance, $\sigma^2_{L_T|IM,NC}$, given by:

$$\mu_{L_T|IM,NC}(im) = \sum_{i=1}^{N_C} \mu_{L_i|IM,NC}(im) \quad (9\text{-}3)$$

$$\sigma^2_{L_T|IM,NC}(im) = \sum_{i=1}^{N_C} \sigma^2_{L_i|IM,NC}(im) + 2\sum_{i=1}^{N_C}\sum_{j=1}^{i-1} \sigma_{L_i,L_j|IM,NC}(im) \quad (9\text{-}4)$$

where $\mu_{L_i|IM,NC}(im)$ and $\sigma^2_{L_i|IM,NC}(im)$ are the mean and variance, respectively, in the loss to component *i* given *IM* = *im*; $\sigma_{L_i,L_j|IM,NC}(im)$ is the covariance in the loss between components *i* and *j* given *IM* = *im*; and N_C is the number of different components in the structure.

In the case of collapse, the total loss is given as the sum of the cost to replace all of the components in the structure (whether they are damaged or not), and also additional costs to account for re-design and demolition [3], making it potentially significantly different from its current market value [13]. The mean, $\mu_{L_T|C}$, and variance, $\sigma^2_{L_T|C}$ in the total loss given collapse are given by:

$$\mu_{L_T|C} = (1 + C_{RDD}) \sum_{i=1}^{N_{CCI}} \mu_{CCI_i} \quad (9\text{-}5)$$

$$\sigma^2_{L_T|C} = (1 + C_{RDD})^2 \left[\sum_{i=1}^{N_{CCI}} \sigma^2_{CCI_i} + 2\sum_{i=1}^{N_{CCI}}\sum_{j=1}^{i-1} \rho_{CCI_i,CCI_j} \sigma_{CCI_i} \sigma_{CCI_j} \right] \quad (9\text{-}6)$$

where μ_{CCI_i} and $\sigma^2_{CCI_i}$ are the mean and variance in the cost of construction cost item (CCI) *i*; N_{CCI} is the number of construction cost items involved in the construction of the structure; ρ_{CCI_i,CCI_j} is the correlation between the cost of construction cost items *i* and *j*; and C_{RDD} is additional costs due to redesign and demolition as a proportion of the total cost. Note the

mean and variance in the loss to repair or replace each component should obviously account for contractor overhead, inflation and location [14].

Equations (9-1)-(9-6) reveal that given the above assumptions, correlations between component losses (which appear in Equations (9-4) and (9-6)) only affect the variance in the total loss, and therefore do not need to be considered when computing only the expected value of the loss [e.g. 1, 2]. As will be shown however, the magnitude of the variance in the total loss is such that, in the author's opinion, it should always be considered when making earthquake risk decisions.

To rationally determine the correlation coefficient in the total loss given no collapse, it is necessary to further examine the covariance in the loss between components *i* and *j*, $\sigma_{L_i,L_j|IM,NC}$. Using the general relationship between covariance and expectations [e.g. 12], $\sigma_{L_i,L_j|IM,NC}$ can be expressed as (where the conditioning on no collapse, *NC*, has been dropped where obvious for brevity):

$$\sigma_{L_i,L_j|IM,NC}(im) = \mu_{L_iL_j|IM}(im) - \mu_{L_i|IM}(im)\mu_{L_j|IM}(im) \tag{9-7}$$

where $\mu_{L_iL_j|IM}(im)$ is the expected value of the product L_iL_j given $IM = im$; and $\mu_{L_i|IM}(im)$ and $\mu_{L_j|IM}(im)$ are the expected losses of components *i* and *j* given $IM = im$, respectively. $\mu_{L_i|IM}$, and similarly $\mu_{L_j|IM}$, are computed by:

$$\mu_{L_i|IM}(im) = \int \mu_{L_i|EDP_i}(edp_i) f_{EDP_i|IM}(edp_i|im) dEDP_i \tag{9-8}$$

where EDP_i is the engineering demand parameter that component *i* is subjected to; $\mu_{L_i|EDP_i}(edp)$ is the expected loss to component *i* given $EDP_i = edp_i$; and $f_{EDP_i|IM}(edp_i|im)$ is the probability density function (pdf) of EDP_i given $IM = im$. Equation (9-8) is an application of the total probability theorem [e.g. 12] and makes the conditional independence assumption that conditioned on EDP_i the mean (and generally the distribution) of L_i is independent of *IM*.

Using the same assumptions in Equation (9-8), the first term on the right-hand side of Equation (9-7), $\mu_{L_iL_j|IM}$, is given by:

$$\mu_{L_iL_j|IM}(im) = \iint \mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j) f_{EDP_i,EDP_j|IM}(edp_i, edp_j|im) dEDP_i dEDP_j \tag{9-9}$$

where $\mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j)$ is the expected value of the product L_iL_j given $EDP_i = edp_i$ and $EDP_j = edp_j$; and $f_{EDP_i,EDP_j|IM}(edp_i, edp_j|im)$ is the bi-variate pdf of EDP_i and EDP_j given $IM = im$.

The expected loss in component i given $EDP_i = edp_i$, $\mu_{L_i|EDP_i}$ can be determined via the use of discrete damage states (DS's) as:

$$\mu_{L_i|EDP_i}(edp_i) = \sum_{k=1}^{N_{DS,i}} \mu_{L_i|DS_k}(ds_k) P_{DS_k|EDP_i}(ds_k|edp_i) \tag{9-10}$$

where $\mu_{L_i|DS_k}(ds_k)$ is the expected loss in component i given $DS_k = ds_k$; $P_{DS_k|EDP_i}(ds_k|edp_i)$ is the probability of $DS_k = ds_k$ given $EDP_i = edp_i$; and $N_{DS,i}$ is the number of damage states for component i. Again, in Equation (9-10) the total probability theorem is used as well as the conditional independence assumption (i.e. that given DS_k, L_i is independent of EDP_i). Details on loss and fragility functions which are needed to determine $\mu_{L_i|DS_k}$ and $P_{DS_k|EDP_i}$ can be found in, for example, Mitrani-Reiser [2] and Porter *et al*. [15].

Similar to Equation (9-10), $\mu_{L_iL_j|EDP_i,EDP_j}$ can be computed by:

$$\begin{aligned}\mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j) &= \sum_{k=1}^{N_{DS,i}} \sum_{l=1}^{N_{DS,j}} \mu_{L_iL_j|DS_k,DS_l}(ds_k, ds_l) P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l|edp_i, edp_j) \\ &= \sum_{k=1}^{N_{DS,i}} \sum_{l=1}^{N_{DS,j}} \left(\mu_{L_i|DS_k}(ds_k)\mu_{L_j|DS_l}(ds_l) + \sigma_{L_i,L_j|DS_k,DS_l}(ds_k, ds_l)\right) x \\ &\qquad P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l|edp_i, edp_j)\end{aligned} \tag{9-11}$$

where $P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l|edp_i, edp_j)$ is the (joint) probability of $DS_k = ds_k$ and $DS_l = ds_l$ in components i and j given $EDP_i = edp_i$ and $EDP_j = edp_j$, respectively; and $\sigma_{L_i,L_j|DS_k,DS_l}(ds_k, ds_l)$ is the covariance in the loss in components i and j given $DS_k = ds_k$ and $DS_l = ds_l$. The term in parentheses on the second line of Equation (9-11) is obtained from the first using an equivalent form of Equation (9-7).

Equations (9-1)-(9-11) completely define those aspects of the seismic loss estimation methodology used here which involve correlations. Other equations which comprise the methodology not involving correlations can be found in Bradley *et al*. [16]. From Equations (9-1)-(9-11), it can be observed that the effects of correlations affect four different terms, namely: $f_{EDP_i,EDP_j|IM}$ (Equation (9-9)), $P_{DS_k,DS_l|EDP_i,EDP_j}$ (Equation (9-11)), $\sigma_{L_i,L_j|DS_k,DS_l}$ (Equation (9-11)), and ρ_{CCI_i,CCI_j} (Equation (9-6)). These four terms are dependent on the correlations in the *EDP*|*IM*, *DS*|*EDP* and *L*|*DS* relationships.

Note that under the adopted framework only a scalar ground motion intensity measure, *IM*, is considered. Should a vector intensity measure be considered [17], then correlations

between the individual *IM* terms should also be accounted for.

9.3.2 Correlations in the *EDP|IM* relationship

The previous section illustrated that the effect of correlations in the *EDP|IM* relationship appears in the joint-distribution, $f_{EDP_i,EDP_j|IM}$. As the marginal ln*EDP|IM* distribution, $f_{\ln EDP_i|IM}$, is typically assumed to have a normal distribution [18], (i.e. $f_{EDP_i|IM}$ has a lognormal distribution) then it is (reasonably) assumed that the joint distribution, $f_{\ln EDP_i,\ln EDP_j|IM}$, is well represented by a bi-variate normal distribution (i.e. $f_{EDP_i,EDP_j|IM}$ has a bi-variate lognormal distribution). Thus $f_{\ln EDP_i,\ln EDP_j|IM}$ is given by:

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}\exp\left[\frac{-1}{2\left(1-\rho_{X,Y}^2\right)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho_{X,Y}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}\right] \tag{9-12}$$

where $X=\ln EDP_i|IM$; $Y=\ln EDP_j|IM$; $\mu_X=\mu_{\ln EDP_i|IM}$ and $\sigma_X=\sigma_{\ln EDP_i|IM}$ are the mean and standard deviation of the ln*EDP*$_i$|*IM* relation; and $\rho_{X,Y}=\rho_{\ln EDP_i|IM,\ln EDP_j|IM}$ is the correlation coefficient between ln*EDP*$_i$|*IM* and ln*EDP*$_j$|*IM*.

9.3.3 Correlations in the *DS|EDP* relationship

Correlations in the *DS|EDP* relationship appear in the joint probability $P_{DS_k,DS_l|EDP_i,EDP_j}$. Determination of this joint probability is complicated by the fact that *DS* is a discrete variable. Aslani [3] and Lee and Kiremidjian [7] propose the use of an optimisation procedure to determine $P_{DS_k,DS_l|EDP_i,EDP_j}$ for a given correlation. Baker [8] notes that this optimisation approach is cumbersome and instead proposes that fragility functions (which define the *DS|EDP* relationship) be considered in terms of the (continuous) EDP which causes the (discrete) DS rather than the probability of DS given EDP. Baker [8] then illustrates how Monte Carlo simulation can be used to determine $P_{DS_k,DS_l|EDP_i,EDP_j}$. Here the approach of considering the fragility function as a continuous function of capacity is taken, but solution via a tractable analytical approach is developed which is significantly more computationally

efficient than simulation-based methods.

Firstly, it is noted that for a single component the damage state k probability, $P_{DS_k|EDP_i}$, is given by the difference in the fragility functions defining the probability of exceeding DS_k and DS_{k+1} [e.g. 19]:

$$\begin{aligned} P_{DS_k|EDP_i}\left(ds_k\middle|edp_i\right) &= F_{DS_k|EDP_i}\left(ds_k\middle|edp_i\right) - F_{DS_{k+1}|EDP_i}\left(ds_{k+1}\middle|edp_i\right) \\ &= F_{DS_k|EDP_i}\left(ds_k\middle|edp_i\right) - \sum_{m=k+1}^{N_{DS,i}} P_{DS_m|EDP_i}\left(ds_m\middle|edp_i\right) \end{aligned} \tag{9-13}$$

where the fragility function, $F_{DS_k|EDP_i}$, is typically assumed to have a lognormal distribution [15]. Use of Equation (9-13) in the case of $k = N_{DS}$ can be handled simply by defining $F_{DS_{N_{DS}+1}|EDP_i} = 0$. The first line of Equation (9-13) is the conventional form by which the damage state probability is defined [e.g. 19], while the equivalent second line is introduced here to aid in the description of the following paragraph. The second line of Equation (9-13) describes the probability of DS_k given EDP_i literally as the probability of DS_k or greater (i.e. $F_{DS_k|EDP_i}$), less all the greater terms (those in the summation).

Using the same logic for the single variable case as in the second line of Equation (9-13), it is trivial to show that the joint probability of DS_k and DS_l given EDP_i and EDP_j is given by:

$$\begin{aligned} P_{DS_k,DS_l|EDP_i,EDP_j}\left(ds_k,ds_l\middle|edp_i,edp_j\right) &= F_{DS_k,DS_l|EDP_i,EDP_j}\left(ds_k,ds_l\middle|edp_i,edp_j\right) \\ &\quad - \sum_{m=k}^{N_{DS,i}} \sum_{\substack{n=l \\ n\neq l \text{ if } m=k}}^{N_{DS,j}} P_{DS_m,DS_n|EDP_i,EDP_j}\left(ds_m,ds_n\middle|edp_i,edp_j\right) \end{aligned} \tag{9-14}$$

Figure 9-1 illustrates schematically the implications of Equation (9-14) for a particular case in which $k = 2$, $l = 2$, $N_{DS,i} = 3$, $N_{DS,j} = 4$. Note that in order to compute Equation (9-14), one must have computed all terms in the summation *a priori*. The numbered square brackets in Figure 9-1 illustrate one possible sequence by which $P_{DS_{k=2},DS_{l=2}|EDP_i,EDP_j}$ can be determined. For example, in step [1], Equation (9-14) becomes: $P_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j} = F_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j}$ (i.e. all the terms in the summation are zero). In step [2], Equation (9-14) becomes: $P_{DS_{k=2},DS_{l=4}|EDP_i,EDP_j} = F_{DS_{k=2},DS_{l=4}|EDP_i,EDP_j} - P_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j}$, where the second term was previously evaluated in step [1]. Thus at each step only $F_{DS_k,DS_l|EDP_i,EDP_j}$ must be computed. It is also important to note that for each component pair (i.e. each i,j pair), Equation (9-11) requires $P_{DS_k,DS_l|EDP_i,EDP_j}$ for all $k = 1\text{-}N_{DS.i}$ and $l = 1\text{-}N_{DS,j}$; thus the sequential process in

Equation (9-14) does not result in any unnecessary computations.

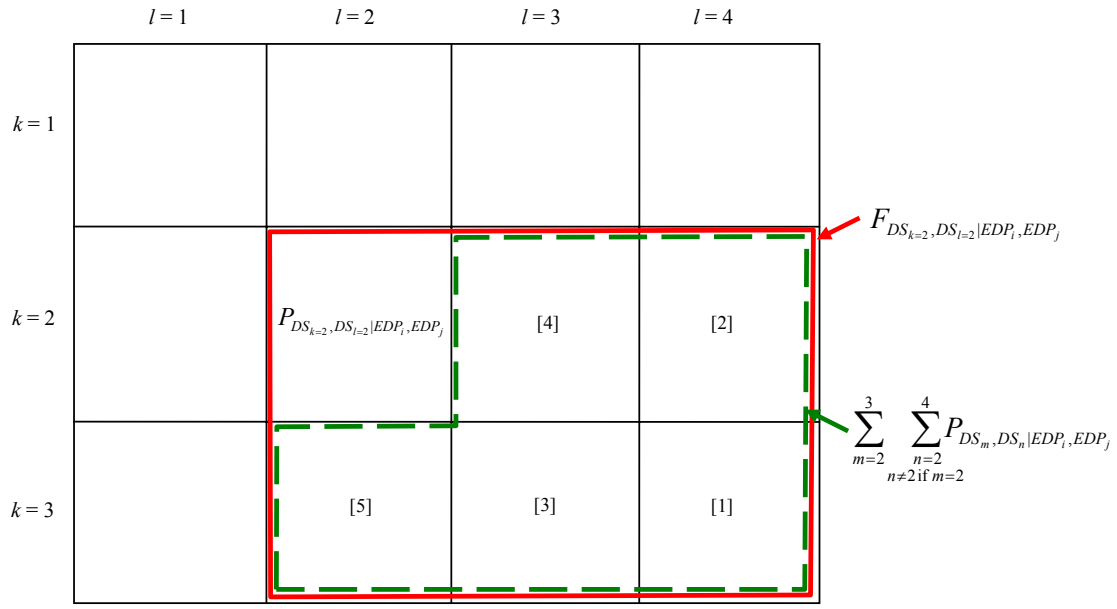


Figure 9-1: Schematic illustration of the computation of damage state probabilities as defined in Equation (9-14) for the case of $k = 2$, $l = 2$, $N_{DS,i} = 3$, $N_{DS,j} = 4$.

In the case of a single component, the damage state fragility function, $F_{DS_k|EDP_i}\left(ds_k|edp_i\right) = P\left(DS_k \geq ds_k|EDP_i = edp_i\right)$, which can be interpreted as the probability of $DS_k \geq ds_k$ given $EDP_i = edp_i$ is equivalent to $P\left(C_{i,k} < EDP_i\right)$, the probability that the demand, EDP_i is greater than the damage state k capacity of component i, $C_{i,k}$ [8]. As $F_{DS_k|EDP_i}$ is typically defined by a lognormal distribution [15] then it follows that $P\left(C_{i,k} < EDP_i\right)$ also has a lognormal distribution, and $P\left(\ln C_{i,k} < \ln EDP_i\right)$ a normal distribution. If it is (reasonably) assumed that $F_{DS_k,DS_l|\ln EDP_i,\ln EDP_j}$ is a cumulative bi-variate normal distribution then it follows by definition that:

$$\begin{aligned} &F_{DS_k,DS_l|\ln EDP_i,\ln EDP_j}\left(ds_k, ds_l|\ln edp_i, \ln edp_j\right) \\ &\quad = P\left(DS_k \geq ds_k . DS_l \geq ds_l|\ln EDP_i = \ln edp_i, \ln EDP_j = \ln edp_j\right) \\ &\quad = P\left(\ln C_{i,k} < \ln edp_i, \ln C_{j,l} < \ln edp_j\right) \\ &\quad = \iint\limits_{\substack{\ln c_{i,k} < \ln EDP_i \\ \ln c_{j,l} < \ln EDP_j}} f_{\ln C_{i,k},\ln C_{j,l}}\left(\ln c_{i,k}, \ln c_{j,l}\right) d\ln C_{i,k}\, d\ln C_{j,l} \end{aligned} \tag{9-15}$$

where $f_{\ln C_{i,k},\ln C_{j,l}}$ is a bi-variate normal distribution pdf (i.e. the same form as that given in

Equation (9-12)) of the component capacities. Thus, only the correlation coefficient, $\rho_{\ln C_{i,k},\ln C_{j,l}}$, defining the correlation between the damage state *k* and *l* capacities of components *i* and *j*, respectively, is required in addition to the conventional fragility function data (which defines the marginal mean and variances).

Equation (9-15) at first may appear to be computationally demanding because of the double integral. However, because of the frequent use of the cumulative bi-variate normal distribution in probability theory, highly efficient numerical algorithms are available. Such algorithms transform the double integral into a single integral and obtain the solution using a combination of analytical integration and as little as 4 Gauss quadrature points [20, 21]. Thus, it can be appreciated that determination of $P_{DS_k,DS_l|EDP_i,EDP_j}$ via Equations (9-14) and (9-15) is orders of magnitude more efficient than the optimization algorithm of Lee and Kiremidjian [7], or solution by Monte Carlo simulation as discussed by Baker [8].

An additional benefit of the formulation given by Equations (9-14) and (9-15) over the optimization approach discussed in Lee and Kiremidjian [7], is that the damage state correlation, $\rho_{\ln C_{i,k},\ln C_{j,l}}$, can potentially be a function of both component types and damage state numbers, while the optimization algorithm allows only a single correlation coefficient per component pair.

9.3.4 Correlations in the *L|DS* relationship

Correlations between the cost to repair damage to different components appear in the term $\sigma_{L_i,L_j|DS_k,DS_l} = \rho_{L_i,L_j|DS_k,DS_l}\sigma_{L_i|DS_k}\sigma_{L_j|DS_l}$ given in Equation (9-11), where $\rho_{L_i,L_j|DS_k,DS_l}$ is the correlation in the loss to components *i* and *j*, due to damage states *k* and *l*, respectively. Correlations in the cost to replace (equivalent to 'repairing' a component in its failure damage state) individual components when collapse occurs is given by ρ_{CCI_i,CCI_j} in Equation (9-6).

9.3.5 Neglected correlations

The previous sections address correlations which appear in the structure-specific seismic loss estimation framework given by Equations (9-1)-(9-10). There are however additional, potentially important, correlations which are not considered because of the assumptions made in the framework. In particular, the conditional independence assumption,

which allows the *L*|*DS*, *DS*|*EDP*, *EDP*|*IM*, and λ_{IM} relationships to be treated independently, and combined using the total probability theorem, means that correlations are only considered within, and not between, these relationships. For example, uncertainty in the capacity of a component of the lateral load resisting system of a structure affects both the damageability of the component itself, but also the dynamic response of the entire structural system. While this uncertainty can be separately considered for the *EDP*|*IM* and *DS*|*EDP* relationships (using a stochastic seismic response model and a component fragility curve, respectively), since they are characterised separately the random capacity for this single component is not correctly accounted for in the strict sense. This is one acknowledged drawback of the conditional independence assumption.

9.4 Causes of and methods to determine correlations

9.4.1 Correlation in the *EDP*|*IM* relationship

Correlations between different EDP's for a given IM (i.e. $\rho_{\ln EDP_i|IM,\ln EDP_j|IM}$) occur due to the dynamic characteristics of the structure and ground motion which it is subjected to. This correlation structure is indeed complex, but can however be determined from the results of multiple time-history or modal pushover [e.g. 22] analyses which are ground-motion dependent. Given a suite of N_{gm} ground motions scaled to a specific value of *IM*, seismic response analyses will yield an N_{edp} x N_{gm} matrix of seismic response, ***EDP*** (i.e. one column for each ground motion and one row for each EDP_i value being monitored), the correlation coefficient for EDP_i and EDP_j can then be computed by:

$$\begin{aligned}\rho_{\ln EDP_i|IM,\ln EDP_j|IM} &= \sum_{k=1}^{N_{gm}}\left(\frac{\ln EDP_{i,k}-\mu_{\ln EDP_i}}{\sigma_{\ln EDP_i}}\right)\left(\frac{\ln EDP_{j,k}-\mu_{\ln EDP_j}}{\sigma_{\ln EDPj}}\right)\\ &= \sum_{k=1}^{N_{gm}}\varepsilon_{i,k}\varepsilon_{j,k}\end{aligned} \tag{9-16}$$

where $EDP_{i,k}=\boldsymbol{EDP}(i,k)$ is the value of i^{th} EDP monitored due to ground motion k; $\mu_{\ln EDP_i}$ and $\sigma_{\ln EDP_i}$ are the mean and standard deviation of lnEDP_i over the N_{gm} different ground motions; and $\varepsilon_{i,k}$ is the so-called standardized residual of lnEDP_i.

The correlations of the N_{edp} different *EDP* values being monitored are defined by a

N_{edp} x N_{edp} symmetric correlation matrix with $0.5\,N_{edp}\,(N_{edp}$ - 1) unique correlation coefficients. Figure 9-2 illustrates the lower-triangular portion of the correlation matrix based on the seismic response analyses of the 10 storey office building discussed in Bradley *et al*. [23, 24]. In Figure 9-2, *EDP* numbers 1-10 are the peak interstorey drift ratios on floors 1-10, and *EDP* numbers 11-21 are peak floor accelerations on the 1st – roof floors. Although the correlation structure is clearly complex, three features can be observed. Firstly, the lower triangular portion of the correlation matrix can be distinguished into three sections: correlations between two peak interstorey drift EDPs (i.e. EDP_i and $EDP_j \leq 10$); correlations between two peak floor accelerations (i.e. EDP_i and $EDP_j \geq 11$) and correlations between a peak interstorey drift and peak floor acceleration (i.e. $EDP_i \leq 10$ and $EDP_j \geq 11$). The magnitude of the correlations in each of these three sections can be put in descending order as: peak floor acceleration vs. peak floor acceleration; peak interstorey drift vs. peak interstorey drift; peak interstorey drift vs. peak floor acceleration. Finally, within any of these three sections, as the value of $\left|EDP_i - EDP_j\right|$ increases, there is a trend for the correlation to reduce.

While the above three observations are insightful, if the correlation matrix can be obtained directly from the results of time-history analyses then such observations are only useful for validation with intuition. Various simplified methods have however been proposed to determine the *EDP*|*IM* relationship for a structure which are ground motion independent [25]. Such ground-motion independent methods therefore do not enable the computation of *EDP*|*IM* correlations as given in Equation (9-16). In the following paragraphs a simple *EDP*|*IM* correlation model is developed for multi-storey buildings based on the structural analysis results of Bradley *et al*. [23, 24] which can be used in conjunction with such ground motion independent simplified methods.

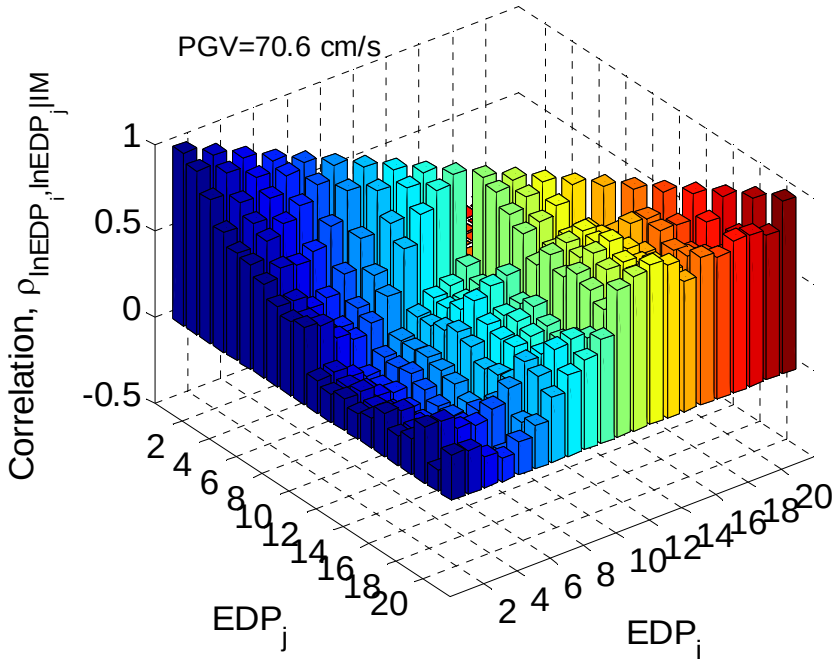


Figure 9-2: Typical correlation matrix of the *EDP|IM* relationship for the case study structure. EDP_1-EDP_{10} are peak interstorey drift ratios on floors 1-10, and EDP_{11}-EDP_{21} are peak floor accelerations on the 1st – roof floors.

Figure 9-3a and Figure 9-3b illustrate the relationship between the correlation coefficient and the number of floors separation, n_{fs}, (equivalent to $\left|EDP_i - EDP_j\right|$) for peak interstorey drifts and peak floor accelerations, respectively. In Figure 9-3a and Figure 9-3b, each point is one correlation coefficient, and each dashed line is the arithmetic mean of the correlations for a specific IM level. The 50 ground motions, 21 EDPs and 9 IM levels used in Bradley *et al*. [23, 24] gives a total of 405 (i.e. 0.5x10x(10 - 1)x9) and 495 correlation coefficients for the peak interstorey drifts and peak floor accelerations, respectively. The similar trends observed in the 9 different arithmetic means in each plot suggests that their is not an overly significant variation in the correlation vs. n_{fs} trend for ground motion intensities resulting in elastic response through to collapse [23, 24]. The solid lines provide simple piecewise linear fits to the data with equation inset in each figure. Figure 9-3c illustrates the correlations coefficient values between a peak interstorey drift and peak floor acceleration, ρ_{a_i,θ_j}, as a function of n_{fs}. It can be seen that the scatter for ρ_{a_i,θ_j} is notably larger than that for ρ_{θ_i,θ_j} and ρ_{a_i,a_j} in Figure 9-3a and Figure 9-3b, suggesting that the correlation coefficient, ρ_{a_i,θ_j}, is significantly dependent other variables in addition to n_{fs}. Despite the larger scatter in Figure 9-3c a simplified value of ρ_{a_i,θ_i} = 0.4 is suggested.

Clearly the above simplified relationships are tentative in the sense that only the seismic response analyses of a single capacity-designed structure have been used in their development. That said, Baker and Cornell [26] also illustrate the drift vs. drift correlation as a function of floor separation for a seven storey structure, and the correlation vs. floor separation trend is remarkably similar to that presented here. This highly simplified method which provides a partial correlation value with no detailed information required is likely to be at a consistent level of accuracy compared to the simplified seismic response analysis used, and is therefore a plausible model until further studies are conducted.

9.4.2 Correlation in the *DS*|*EDP* relationship

Correlation in the *DS*|*EDP* relationships of different components occurs due to the shared uncertainties in the EDP values at which the specific DS's occur. This uncertainty can be separated into shared uncertainty due to the EDP definition and the component capacity.

While the conditional independence assumption is made in the loss assessment framework for simplicity and tractability (i.e. that given *EDP* the damage state probability is independent of *IM*), common EDP's, such as peak interstorey drift and peak floor acceleration do not account for the frequency content and duration of the seismic excitation. Thus, for example, a large magnitude earthquake (with a correspondingly long duration) is likely to lead to more damage throughout the structure than a small magnitude earthquake (with short duration) which causes the same (peak) demand vector, ***EDP*** = ***edp***, but a different response history [27] (i.e. the EDP used is an *insufficient* [28] predictor of DS).

Shared uncertainty in the capacity of the different components is present due to the dependence in the material properties which the components are comprised of, and the similarity in the installation techniques required.

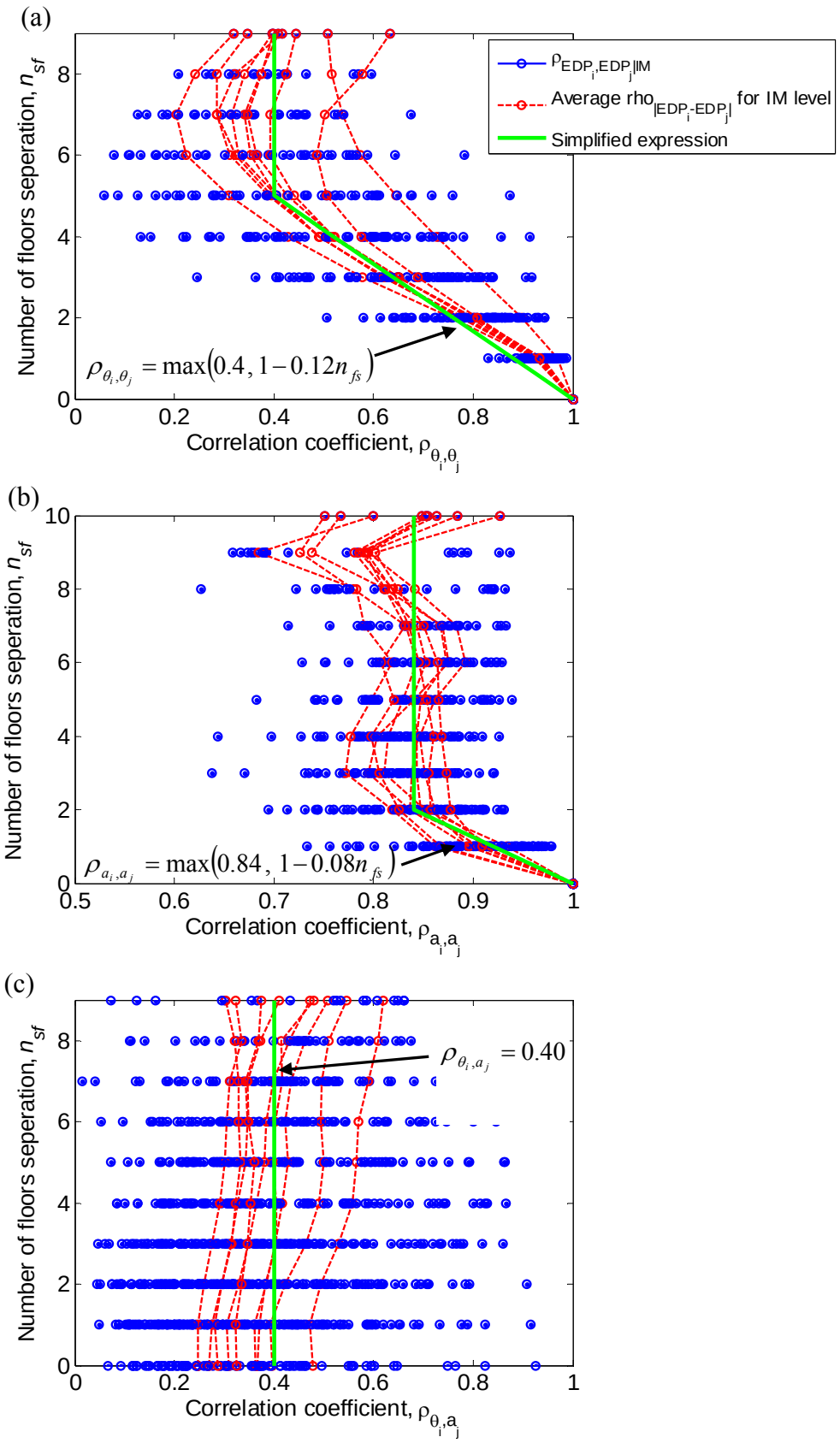


Figure 9-3: Trends in the EDP|IM correlations as a function of the number of floors separation of the demand being measured: (a) correlations between peak interstorey drifts on different floors; (b) correlations between peak floor accelerations on different floors; and (c) correlations between peak interstorey drift and peak floor acceleration on the different floors.

Unlike the correlations in the *EDP*|*IM* relationship, which, for a specific structure, can be obtained from the results of ground-motion dependent numerical simulations, there is a paucity of data for use in determining correlations between *DS*|*EDP* relations. Lee and Kiremidjian [7] and Aslani [3] both discuss the consideration of *DS*|*EDP* correlations but do not present any means or data to determine the correlations. Baker and Cornell [9] propose the use of the generalised equi-correlated model [e.g. 29] for determining the correlation in a 'collapsed' *L*|*EDP* relationship. Here the generalised equi-correlated model is developed for the *DS*|*EDP* relationship (and also the *L*|*DS* relationship in the next section). Baker [8] notes that there are several methods by which correlations can be (in future) determined from observations. It is suggested therefore that a generalised equi-correlated model is used to initially determine correlations, and as observational data becomes available in the future they can be used in a Bayesian framework to update specific correlation values.

The generalised equi-correlated model adopted for the *DS*|*EDP* relationship is composed of five mutually independent variables. These five variables have been chosen to account for the main contributing sources of uncertainty, and can be easily handled in the loss assessment computation (i.e. a more robust model would further break down some of these variables, but the input and computational requirements to compute the correlation would be significantly increased). The five variables considered represent: (i) uncertainty due to the EDP definition common to all components (e.g. the aforementioned example relating to duration of shaking); (ii) uncertainty due to the EDP definition common to components subjected to the same EDP (e.g. two components both sensitive to 2nd floor acceleration); (iii) uncertainty in seismic capacity common to components made of the same material; (e.g. a structural beam and structural wall both made of concrete); (iv) uncertainty in seismic capacity common to components of the same type (e.g. two different concrete structural beams); and (v) uncertainty in seismic capacity of a single component independent of all other components).

Based on the five variables explained above, the total uncertainty in the *DS*|*EDP* relationship for component *i*, which is dependent on EDP_k is given by:

$$\sigma^2_{\ln EDP_{k_i}|DS} = \left(\sigma^2_{structure_i} + \sigma^2_{EDP_{k_i}}\right) + \left(\sigma^2_{mat_i} + \sigma^2_{comptype_i} + \sigma^2_{comp_i}\right) \tag{9-17}$$

where $\sigma^2_{structure_i}$ and $\sigma^2_{EDP_{k_i}}$ are the sheared uncertainties due to the EDP definition for the entire structure, and EDP_k, respectively for component *i* (due to insufficiency of the EDP's used); $\sigma^2_{mat_i}$, $\sigma^2_{comptype_i}$, and $\sigma^2_{comp_i}$ are the shared uncertainties in the capacity for material type,

component type, and component, respectively, for component *i*. It is assumed that each of the five variables is either independent or perfectly correlated with the same variable of a different component.

It should be noted that some fragility functions have been developed without consideration of the insufficiency of the EDP, typically those based on quasi-static laboratory experiments [30]. In such cases, the total uncertainty represents only capacity uncertainty. Further research is needed to more accurately determine the magnitude of the uncertainty due to using insufficient EDP's. Alternatively, using EDP's which account for intensity, frequency content and duration of excitation can be expected to significantly reduce the demand uncertainty portion of the fragility function uncertainty.

A limitation of the equi-correlated model as adopted in this application is that no consideration has been made of the particular damage states of the components. For example, the dependence between cracking damage states in a concrete wall and concrete beam (both dependent on the concrete tensile strength) will have a higher correlation that cracking in the concrete wall and failure in the concrete beam (failure in the ductile beam being primarily dependent on the reinforcing steel properties). As no empirical data is currently available to warrant a damage state dependent correlation such an effort is not pursued here, although it is again noted that mathematically speaking a damage state dependent correlation is not a problem for the framework presented in this chapter.

From Equation (9-17), and using the relation, $Cov[a,b+c]=Cov[a,b]+Cov[a,c]$, the covariance between two different components, *i* and *j*, is given by:

$$\begin{aligned} Cov\left[\ln EDP_{k_i}\,|DS, \ln EDP_{k_j}\,|DS\right] &= \left(\sigma_{structure_i}\sigma_{structure_j} + \delta_{k_i k_j}\sigma_{EDP_{k_i}}\sigma_{EDP_{k_j}}\right) \\ &+ \left(\delta_{mat_i mat_j}\sigma_{mat_i}\sigma_{mat_j} + \delta_{comptype_i comptype_j}\sigma_{comptype_i}\sigma_{comptype_j} + \delta_{ij}\sigma_{comp_i}\sigma_{comp_j}\right) \end{aligned} \quad (9\text{-}18)$$

where δ_{ij} is the Kronecker delta function equal to one if $i=j$ and zero otherwise (i.e. corresponding to zero of perfect correlations). Note that no Kronecker delta function is necessary for the product $\sigma_{structure_i}\sigma_{structure_j}$, since it is common to all components and thus the Kronecker delta function would always equal one.

Based on the author's judgement, of the total variance in the damage state uncertainty (Equation (9-17)): 30% is assumed to be due to seismic demand uncertainty and 70% due to component capacity uncertainty; 67% of the demand uncertainty is assumed to be common to the entire structure ($\sigma^2_{structure_i}$), and 33% to a specific EDP ($\sigma^2_{EDP_k}$); 50% of the capacity uncertainty is assumed to be common to specific material types ($\sigma^2_{mat_i}$), 35% common to

specific component types ($\sigma^2_{comptype_i}$) and 15% specific to each component ($\sigma^2_{comp_i}$). If, in Equation (9-18), each of the five variables are written as a proportion of the total uncertainty (e.g. $\sigma^2_{structure_i} = 0.3\left(0.67\sigma^2_{\ln EDP_{k_i}|DS}\right)$) then the correlation coefficient, becomes, after simplification:

$$\rho_{\ln EDP_{k_i}|DS,\ln EDP_{k_j}|DS} = 0.3\left(0.67 + 0.33\delta_{k_i k_j}\right) + 0.7\left(0.5\delta_{mat_i mat_j} + 0.35\delta_{comptype_i comptype_j} + 0.15\delta_{ij}\right) \quad (9\text{-}19)$$

As illustrated by Equation (9-19), the assumption that the standard deviation values for each of the five variables are a function of the total uncertainty in the fragility function, as opposed to $\sigma^2_{structure_i}$ being the same for all components, means that the correlation is in fact dependent only on the four δ_{ij} terms. The argument against use of constant values for the variables is based on the likelihood that different components will be less or more sensitive to different variables. The fact that some fragility functions have (logarithmic) standard deviations as low as 0.28 considering both demand and capacity portions of the uncertainty [e.g. 31], while others can be in the vicinity of 0.6 without considering the demand portion of the uncertainty [e.g. 19] suggests that the magnitude of the variable uncertainties is not likely to be the same for different components.

9.4.3 Correlation in the *L*|*DS* relationship

Correlations between the *L*|*DS* relationships of different components arise due to similarity in the repair actions required to repair the component. This similarity will dictate, for example, whether the same labourers repair both components *i* and *j*, and whether the same materials are required to conduct the repair. Both of these effects will directly influence the correlation in the loss (be it direct repair cost or time taken to perform the repairs).

While the consideration of such correlations for seismic loss estimation is a relatively new problem, correlations between construction costs for determining total project budgets and cost contingencies have received attention in engineering management literature [e.g. 32, 33]. Not surprisingly, the same problem of limited empirical data to determine such correlations is the central issue. Touran and Wiser [34] provide correlation coefficients between unit costs of general construction items such as concrete, metals, electrical, etc. based on data from 26 projects, which Aslani [3] used in seismic loss estimation. Apart from a correlation of 0.79 between electrical and mechanical costs, all correlation values in Touran

and Wiser [34] are below 0.51, most likely due to the broad construction category definitions used. An alternative approach adopted by Hudak and Maxwell [33] is to use a so-called macro approach in which common external (or macro-) factors are identified which are common to multiple cost items. The macro approach of Hudak and Maxwell [33] is similar to the aforementioned generalised equi-correlated model used for the *DS|EDP* correlations, but allows the use of coefficients for each of the macro factors, and also partial correlation between the same macro-factors in different components. In comparison, coefficients of one, and either none or perfect correlations between the variables were used in the generalised equi-correlated model used for the *DS|EDP* correlations. While the model of Hudak and Maxwell [33] is more general than the model used herein, it also requires additional data (i.e. values of the coefficients and partial correlations). Thus, such a model may be more appropriate in future, when additional data warrant such generalisation.

The generalised equi-correlated model adopted for the L|DS relationship is composed of three variables, which represent: (i) uncertainty in repair cost/duration common to components made of the same material; (e.g. a structural beam and structural wall both made of concrete); (ii) uncertainty in repair cost/duration common to components of the same type (e.g. two different concrete structural beams); and (iii) uncertainty in repair cost/duration of a single component independent of all other components). Based on these three variables, the total uncertainty in the L|DS relationship for component *i*, given some DS is given by:

$$\sigma^2_{\ln L_i|DS} = \sigma^2_{mat_i} + \sigma^2_{comptype_i} + \sigma^2_{comp_i} \tag{9-20}$$

where $\sigma^2_{mat_i}$, $\sigma^2_{comptype_i}$, and $\sigma^2_{comp_i}$ are the uncertainty in the repair cost/duration for material type, component type, and component, respectively, for component *i*. The covariance in the repair cost/duration between two different components, *i* and *j*, is therefore given by:

$$\begin{aligned} Cov\left[\ln L_i|DS, \ln L_j|DS\right] = {} & \delta_{mat_i mat_j}\sigma_{mat_i}\sigma_{mat_j} + \delta_{comptype_i comptype_j}\sigma_{comptype_i}\sigma_{comptype_j} \\ & + \delta_{ij}\sigma_{comp_i}\sigma_{comp_j} \end{aligned} \tag{9-21}$$

As with the equi-correlated model for the *DS|EDP* relationship, the *L|DS* equi-correlated model does not consider particular damage states of the components, and only the components themselves. Based on the author's judgement, of the total variance in the damage state uncertainty (Equation (9-20)), 50% is assumed to be common to specific material types ($\sigma^2_{mat_i}$), 35% common to specific component types ($\sigma^2_{comptype_i}$), and 15% specific to each component ($\sigma^2_{comp_i}$). Based on these assumptions, the correlation coefficient for the *L|DS* equi-correlated model is:

$$\rho_{\ln L_i|DS,\ln L_j|DS} = 0.5\delta_{mat_i mat_j} + 0.35\delta_{comptype_i comptype_j} + 0.15\delta_{ij} \quad (9\text{-}22)$$

While the adopted *L*|*DS* equi-correlated model does not consider partial correlations in each of the variables, a clear benefit of using Equation (9-22), as opposed to the construction correlation data of Touran and Wiser [34] is that different components made of a similar material are not automatically assumed to be perfectly correlated.

9.5 Correlations and epistemic uncertainty

Correlations and epistemic uncertainties are coupled in two ways in a rigorous seismic loss assessment. Firstly, because of the various assumptions made in determining the correlation coefficients, these values have some associated epistemic (knowledge-based) uncertainty. Secondly, epistemic uncertainties in the input values in a seismic loss assessment tend to be correlated at different values of the dependent variables. Neglect of both of these two points can potentially lead to erroneous decision making and a brief discussion is given below concerning these two points.

9.5.1 Epistemic uncertainty in correlation coefficients

There is potentially significant epistemic (knowledge) uncertainty in the correlation coefficients examined to date in this chapter due to the equi-correlated model assumed and judgement required in determining the magnitude of the variables (in the case of the *DS*|*EDP* and *L*|*DS* relationships) and assumptions made in seismic response modelling (in the case of the *EDP*|*IM* relationship). The method by which epistemic uncertainty in correlation coefficients is considered in seismic loss estimation will be partially dependent on how epistemic uncertainty is considered in each of the *EDP*|*IM*, *DS*|*EDP* and *L*|*DS* relationships. Here, for brevity, only a single possible approach is considered which is in line with developments to date, and is, the author's opinion, a good compromise between simplicity and accuracy.

Epistemic uncertainties in the *EDP*|*IM* relationship are treated in a discrete fashion via the use of logic-trees, much the same as those used for treating epistemic uncertainties in seismic hazard analyses [35]. In this discrete case, the epistemic uncertainty in the correlation coefficients between various EDP's is simply accounted for by computing different correlation coefficients for each set of analyses performed using the different logic tree

models. This discrete logic-tree approach is preferred over a continuous treatment of the epistemic uncertainty because of the anticipated complexity in defining the epistemic uncertainties and their correlation with the other EDP epistemic uncertainties in continuous form.

Epistemic uncertainties in both the *DS|EDP* and *L|DS* relationships are treated in continuous form by randomly generating the distribution parameters (mean and variance) of the fragility and loss functions [e.g. 30]. Epistemic uncertainty in the values of the correlation coefficients in the *DS|EDP* and *L|DS* relationships can be considered by having uncertain variable magnitudes in the generalised equi-correlated model. Hence, for each realization of the magnitude of the variables a correlation coefficient can be computed. This approach is both simple and also ensures that the correlation matrix is strictly semi-positive definite. If the *DS|EDP* or *L|DS* correlation matrix is required for the simulation of correlated random numbers in the loss estimation [e.g. 6, 14] then it is possible that the Cholesky decomposition of a non semi-positive definite correlation matrix may contain imaginary numbers.

A uniform probability distribution is assumed for each of the variables in the generalised equi-correlated model. The parameters of the uniform distribution (the lower and upper bounds), were defined as 0.5 and 1.5 times (i.e. ±50% of) the judgement-based value. For example, in the *DS|EDP* relationship, $\sigma^2_{structure_i}$ is 0.3*0.67=0.201 of the total uncertainty, and thus its uniform distribution has lower and upper bounds of 0.101 and 0.302, respectively. Figure 9-4 illustrates the distribution of the *DS|EDP* correlation coefficients for two cases based on 10000 Monte Carlo simulations of the random magnitudes of the equi-correlated model variables. It is noted that despite the use of the uniform distribution for the variables: (i) the distribution of the correlation coefficient in Figure 9-4a is approximately normal (although doubly truncated); (ii) Figure 9-4b is skewed to the left; and (iii) the variance in Figure 9-4a is larger than in Figure 9-4b. All of the above three observations are inline with the Fisher transformation of the correlation coefficient having a normal distribution [36].

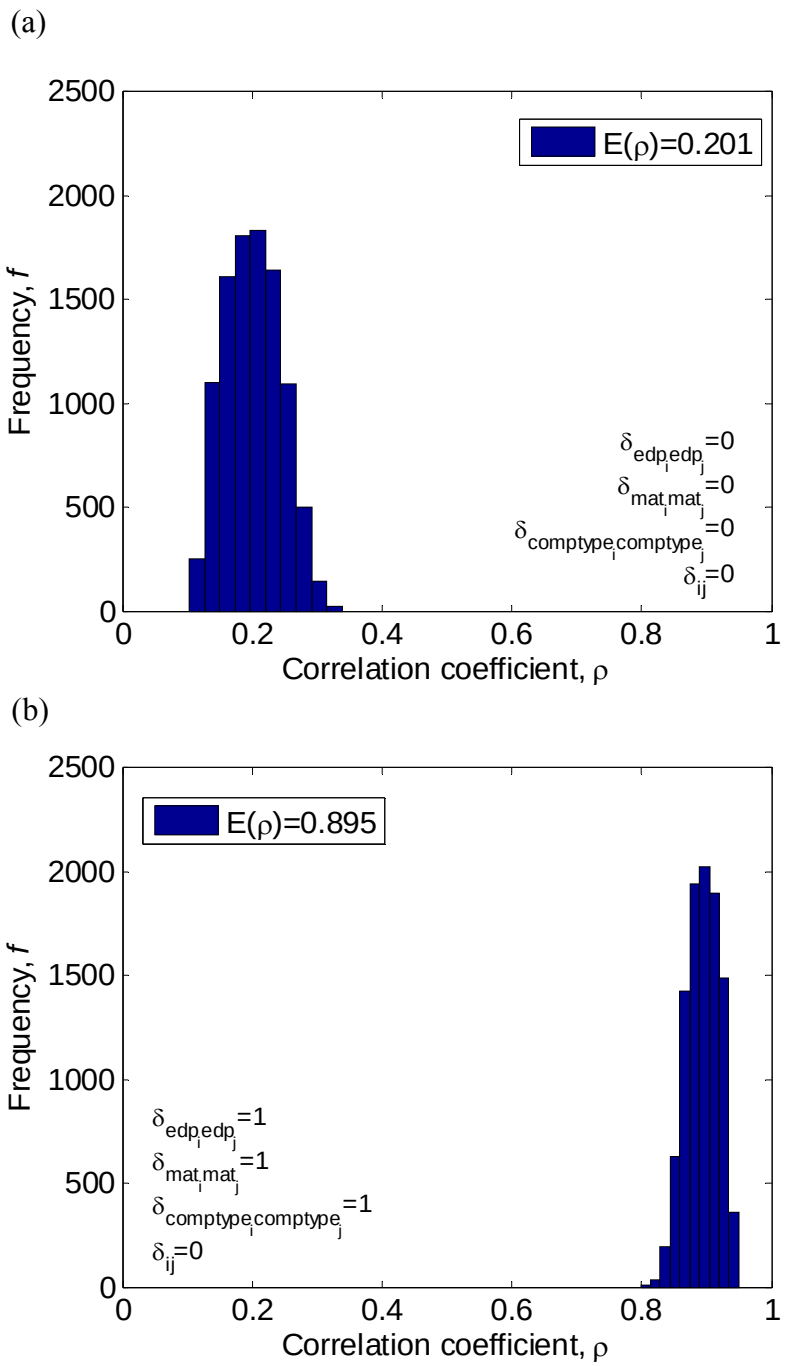


Figure 9-4: Distribution of the correlation coefficient based on 10000 Monte Carlo simulations with each variable having a uniform distribution with a variation of ±50% of its best-estimate value.

9.5.2 Correlations in epistemic uncertainties

In addition to the aleatory uncertainty correlations between different variables which have been discussed in previous sections, there are also correlations between epistemic uncertainties which should be considered in seismic loss estimations.

Bradley [37] discusses the typical correlation structure of epistemic uncertainty in

seismic hazard curves. If a logic tree approach [35] is adopted then epistemic uncertainty correlations are implicitly accounted for, while equations are also available if epistemic uncertainties are treated in a parametric form.

Epistemic uncertainties in the *EDP|IM* relationship resulting from modelling uncertainties in seismic response analysis are starting to gain attention in literature [e.g. 38, 39]. Because of the likely dependence on model configuration it is suggested that correlation between *EDP|IM* epistemic uncertainties is also handled in a non-parametric logic tree format.

Bradley [30] discussed causes of and methods to determine epistemic uncertainties and their correlations in component fragility functions. As noted by Aslani [3], there is currently no literature suggesting the magnitude of epistemic uncertainty (and their correlations) in component loss functions. Clearly, future work is required in this area, as it is expected the magnitude of the epistemic uncertainty will be significant.

9.6 Case-study seismic loss estimation results

This section illustrates the effects of five different assumptions regarding the correlation coefficients discussed to date. The seismic loss estimation results presented are for a typical New Zealand 10 storey office building, and include losses resulting from damage of structural, non-structural, and contents components. Details on the seismic hazard, input ground motions and seismic response analyses can be found in Bradley *et al*. [23], while component inventory data is given in Bradley *et al*. [24]. The five different assumptions regarding correlations are: (i) all correlations zero; (ii) all correlations perfect; (iii) all best-estimate partial correlation; (iv) *EDP|IM* correlations perfect and *DS|EDP* and *L|DS* correlations zero; (v) *EDP|IM* correlations perfect and *DS|EDP* and *L|DS* best-estimate partial correlation. Assumptions (i) and (ii) represent the bounding solutions, while assumptions (iv) and (v) represent attempts to approximate the 'correct' assumption (iii) which are less computationally demanding. For simplicity, epistemic uncertainties in the estimated correlation coefficients are not considered in the example to follow.

9.6.1 Total loss given collapse, *L|C*

Figure 9-5 illustrates the probability density function (pdf) for the total loss given collapse, for three different *L|DS* correlation assumptions (assumptions (i)-(v) above result in only three unique solutions in this case). As previously noted, in the adopted framework

correlation assumptions do not affect the expected collapse loss (Equation (9-5)) which was $12 million, and the cost distribution is assumed to be lognormal [34]with variance obtained from Equation (9-6). Figure 9-5 illustrates that the lognormal standard deviation assuming perfect correlations is three times larger than assuming no correlations. Using partial correlations yields a standard deviation which is 24% larger than that assuming no correlation assumption, similar to that given by Aslani [3].

Comparison of the magnitude of the lognormal standard deviations for the total loss given collapse indicates that the results presented here are somewhat smaller than those given by Aslani [3] (which were lognormal standard deviations of 0.24, 0.4, and 0.65 for none, partial, and perfect correlations, respectively). This is largely due to significantly larger number of components and component types considered compared with Aslani [3], and as Aslani [3] illustrates, the lognormal standard deviation will reduce as the number of partially correlated components increases. This same logic applies to the lognormal standard deviations for the *L|IM,NC* and *L|IM* relationships discussed below.

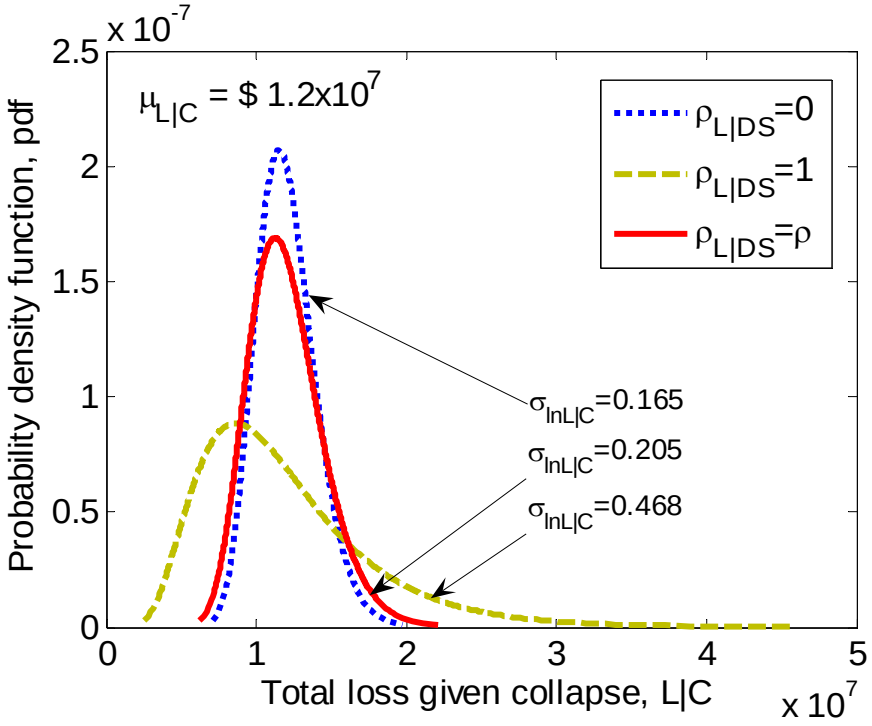


Figure 9-5: Distribution of total loss given structural collapse with various correlation assumptions.

9.6.2 Total loss given no collapse, *L|IM,NC* and total loss, *L|IM*

Figure 9-6a and Figure 9-6b illustrate the lognormal standard deviation in the total loss given no collapse and the total loss, respectively, while Figure 9-6c and Figure 9-6d give the

corresponding error ratios compared to the case of all correlations allowed to be partial. Firstly, it is noted from Figure 9-6a and Figure 9-6c that the bounding assumptions of none and perfect correlations are typically in error by over 50%, and the error is relatively constant as a function of the ground motion IM (which is peak ground velocity, *PGV*). Secondly, both approximations (iv) and (v) result in standard deviations which are in close agreement with the best-estimate solution.

The standard deviation values in Figure 9-6b are similar those in Figure 9-6a for small *PGV*, but tend to the standard deviation given collapse (Figure 9-5), for increasing *PGV* values as indicated by Equation (9-2). The local maxima in the lognormal standard deviation of the total loss, $\sigma_{\ln L|IM}$, for the case of no correlations occurs due to the contribution of the third term on the right hand side of Equation (9-2). This term is most significant at the *PGV* value for which $\left(1-P_{C|IM}\right)P_{C|IM}$ is maximised (i.e. $P_{C|IM}=0.5$). The effect of this third term for the other four correlation assumptions is less pronounced because in these cases $\sigma_{\ln L|C}$ is significantly less than $\sigma_{\ln L|IM,NC}$. Figure 9-6d illustrates that, for this particular structure, the upper bound assumption of perfect correlations over-approximates $\sigma_{\ln L|IM}$ by about 50% over a wide range of *PGV*, while the error associated with the assumption of no correlations is also up to 50% for small *PGV* values, and decreases with increasing *PGV*.

Considering the magnitude of the lognormal standard deviation in the case of partial correlations it is inappropriate to use the central limit theorem to assume that $\sigma_{\ln L|IM,NC}$ or $\sigma_{\ln L|IM}$ have a normal distribution, as done so by Aslani [3]. For a lognormal standard deviation of 0.8, for example, there is a probability of 0.15 of having a loss less than zero if a normal distribution is assumed. It is therefore assumed that the *L|IM,NC* and *L|IM* distributions are lognormal based on the results of Krawinkler [40] and similar Monte Carlo analyses conducted by the authors.

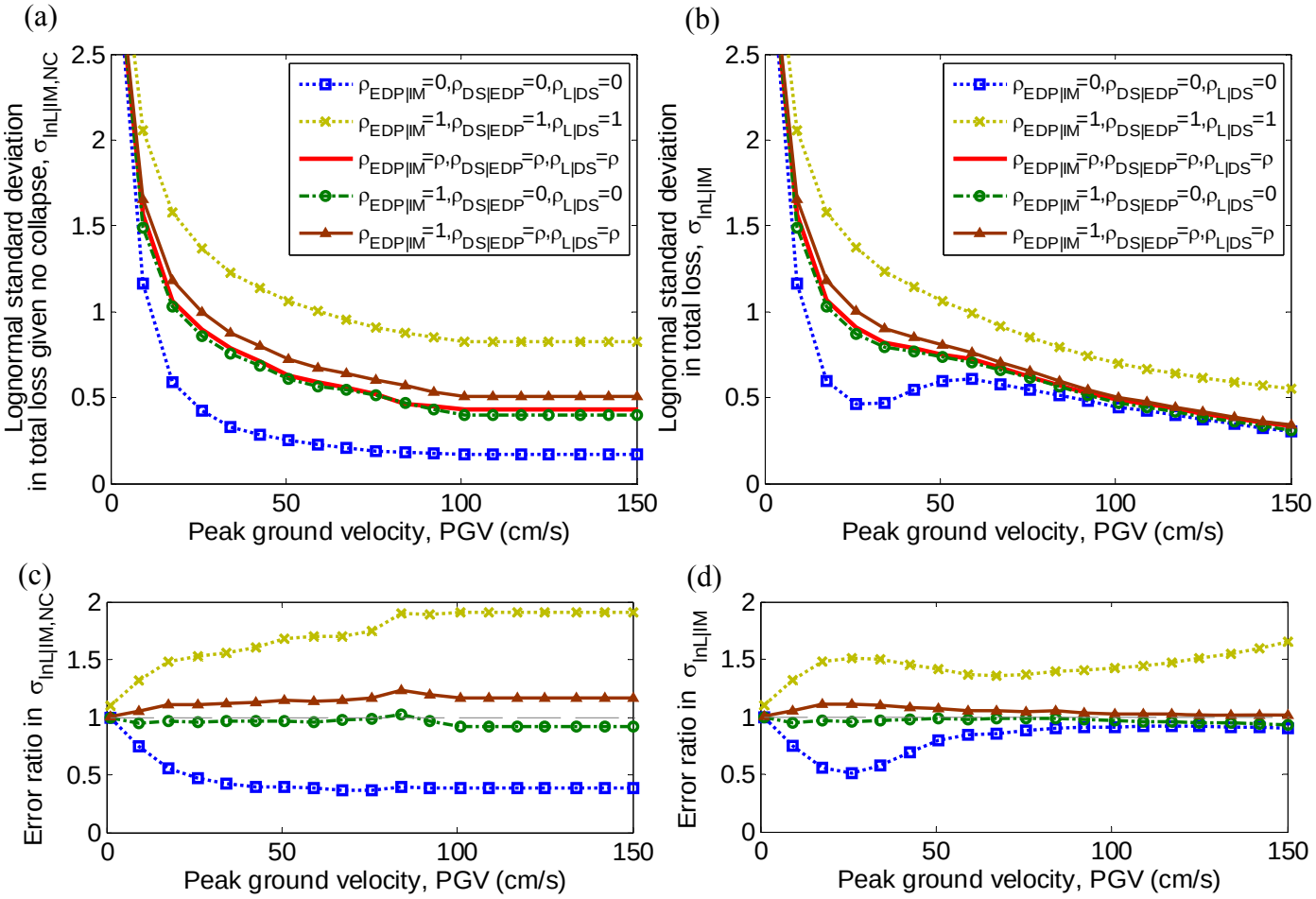


Figure 9-6: Effect of correlation assumptions on the dispersion in the total loss: (a) lognormal standard deviation in the total loss given no collapse; (b) lognormal standard deviation in total loss; (c) error ratios in the dispersion in the total loss given no collapse; and (d) error ratios in the dispersion in the total loss.

9.6.3 Loss hazard, P_L.

Figure 9-7a illustrates the loss hazard curve for the case study structure based on the five different correlation approximations, while Figure 9-7b and Figure 9-7c illustrate the error ratios in loss and annual probability of exceedance, respectively. Mathematical details of the loss hazard computation can be found in References [3, 16]. Figure 9-7b illustrates that for exceedance probabilities greater than 0.01 the bounding solutions of zero and perfect correlations yield errors in the loss of around 50%. For this same P_L range, which correspond to losses less than $1x10^6$ (i.e. \$1 million), the error for the zero correlations assumption is larger than 20%, while for the perfect correlations assumption the error is up to 20%. For losses in excess of $2x10^6$ (with corresponding probabilities of exceedance less than $2x10^{-3}$), the trends are reversed with the zero correlations assumption giving an under-prediction and the perfect correlation assumption giving an over-prediction. For losses larger than the mean

value of the collapse loss (i.e. 1.2×10^7) the loss hazard curve is particularly sensitive to the assumption of the *L|DS* correlation with error ratios greater than 2 when perfect *L|DS* correlations were assumed and less than 0.5 when zero *L|DS* correlations were assumed.

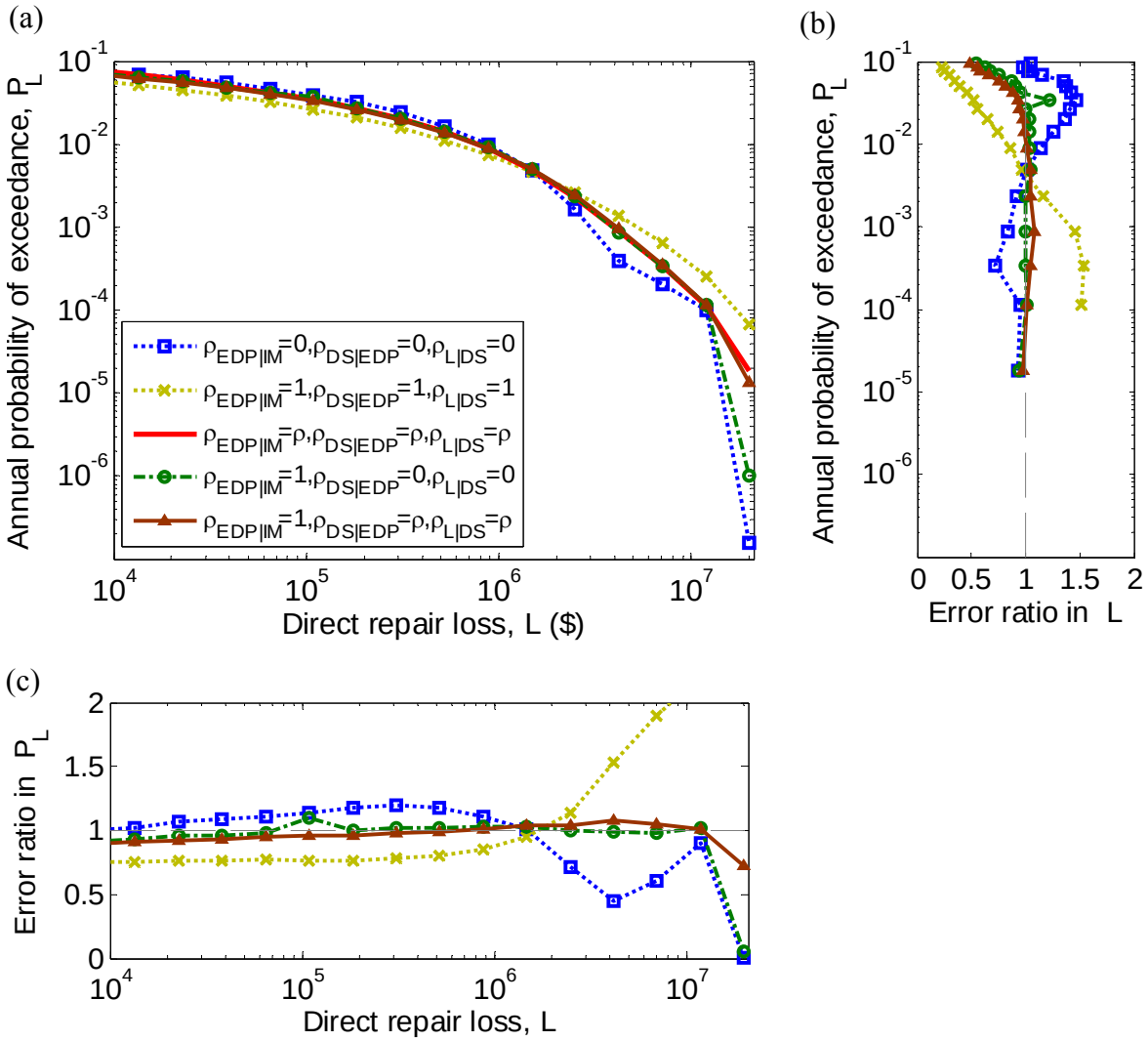


Figure 9-7: (a) Resulting loss hazard curves based on different assumptions regarding correlations; (b) error ratio in loss, *L*; and (c) error ratio in annual probability of exceedance, P_L.

9.6.4 Computational demand

Table 9-1 presents the computational times required when in performing the seismic loss assessment on a Pentium 4 processor with 3.0 GHz CPU and 512 MB RAM using the seismic loss assessment tool (SLAT) [41], which utilizes the magnitude-oriented adaptive quadrature algorithm [42]. As discussed by Bradley and Lee [10], and evident in Table 9-1, the effect of non-zero correlations drastically increases the computational demand to perform the analysis. Table 9-1 illustrates that the loss hazard involves approximately 20 times more

computational time than the *L*|*IM* relationships, and it can be seen that the using perfect and partial correlations requires approximately 50-times and 3200-times more computational time, respectively, than the assumption of no-correlations. The fourth and fifth rows of Table 9-1 which correspond to correlation approximations (iv) and (v), respectively, illustrate that by assuming perfect *EDP*|*IM* correlations and either zero or partial correlations for the *DS*|*EDP* and *L*|*DS* relationships, the computational time to run the analysis is significantly reduced compared to assuming all correlations are partial. The small difference in computational times for correlation assumptions (iv) and (v) illustrate that the additional computational demand due to the *DS*|*EDP* and *L*|*DS* correlations is minimal compared to considering *EDP*|*IM* correlations. By assuming perfect *EDP*|*IM* correlations the double-integral in Equation (9-9) reduces to a single integral as discussed by Bradley and Lee [10], which significantly reduces the computational time. Since correlation approximation (v) uses perfect *EDP*|*IM* correlations and partial *DS*|*EDP* and *L*|*DS* correlations then it will always give a larger value for the lognormal standard deviation in the total loss (i.e. both $\sigma_{\ln L|IM,NC}$ and $\sigma_{\ln L|IM}$), than using all partial correlations. Thus, based on the accuracy of correlation assumption (iv), it may be viewed as a good approximation if it is deemed that the computational demand associated with assuming all partial correlations is excessive.

Table 9-1: Computational times for seismic loss analyses.

Correlation assumption*		Loss vs. *IM*	Loss vs. P_L
(i)	$\rho_{EDP\|IM}=0;\ \rho_{DS\|EDP}=0;\ \rho_{L\|DS}=0$	2.6 sec	52 sec
(ii)	$\rho_{EDP\|IM}=1;\ \rho_{DS\|EDP}=1;\ \rho_{L\|DS}=1$	130 sec (2.1 min)	2500 sec (41 min)
(iii)	$\rho_{EDP\|IM}=\rho;\ \rho_{DS\|EDP}=\rho;\ \rho_{L\|DS}=\rho$	8200 sec (2.3 hr)	180000 sec (49 hr)
(iv)	$\rho_{EDP\|IM}=1;\ \rho_{DS\|EDP}=0;\ \rho_{L\|DS}=0$	130 sec (2.1 min)	2800 sec (46 min)
(v)	$\rho_{EDP\|IM}=1;\ \rho_{DS\|EDP}=\rho;\ \rho_{L\|DS}=\rho$	190 sec (3.2 min)	3700 sec (1.0 hr)

*Correlation coefficient values of 0, 1, and ρ correspond to zero, perfect, and partial correlations, respectively.

9.7 Conclusions

The consideration of component correlations in seismic loss estimation has been limited by methodological tractability, increased computational demand, and a paucity of data for their computation. This chapter has presented a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered. Methods to

determine the necessary correlations were discussed, particularly those which can be used in the absence of sufficient empirical data and rely somewhat on judgement. The effects of various assumptions regarding correlations were illustrated via application to a case-study office structure. It was observed that certain correlation assumptions can lead to errors in excess of 50% in the lognormal standard deviation in the loss given intensity and loss hazard relationships, while full consideration of partial correlations requires in excess of 50-times more computational time than other correlation assumptions.

9.8 Acknowledgements

Financial support of the first author from the New Zealand Tertiary Education Commission Bright Futures scheme is appreciated.

9.9 References

[1] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1973-1997, DOI: 10.1002/eqe.694.

[2] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[3] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[4] Porter KA and Kiremidjian AS. Assembly-based vulnerability and its uses in seismic performance evaluation and risk-management decision-making. John A.Blume Earthquake Engineering Center, Stanford, CA, 2001. 216.

[5] Iervolino I, Giorgio M, and Manfredi G. Expected loss-based alarm threshold set for earthquake early warning systems. *Earthquake Engineering and Structural Dynamics* 2007; **36**(9): 1151-1168, DOI: 10.1002/eqe.675.

[6] ATC-58. Guidelines for seismic performance assessment of buildings: ATC-58 35% Draft. Applied Technology Council, 2007. http://www.atcouncil.org/atc-58.shtml (last accessed 19/11/08)

[7] Lee R and Kiremidjian AS. Uncertainty and correlation for loss assessment of spatially distributed systems. *Earthquake Spectra* 2007; **23**(4): 753-770, DOI: 10.1193/1.2791001.

[8] Baker JW. Introducing correlation among fragility functions for multiple components, in *14th World Conference on Earthquake Engineering*, Beijing, China, 2008, 8.

[9] Baker JW and Cornell CA. Uncertainty propagation in probabilistic seismic loss estimation. *Structural Safety* 2008; **30**(3): 236-252.

[10] Bradley BA and Lee DS. Accuracy of approximate methods of uncertainty propagation in loss estimation. *Structural Safety* 2009: (in press), DOI: 10.1016/j.strusafe.2009.04.001.

[11] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).

[12] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[13] Kircher CA, Reitherman RK, Whitman RV, and Arnold C. Estimation of earthquake losses to buildings. *Earthquake Spectra* 1997; **13**(4): 703-720.

[14] Porter KA, Kiremidjian AS, and LeGrue JS. Assembly-Based Vulnerability for Buildings and Its Use in Performance Evaluation. *Earthquake Spectra* 2001; **17**(2): 291-312.

[15] Porter K, Kennedy R, and Bachman R. Creating Fragility Functions for Performance-based Earthquake Engineering. *Earthquake Spectra* 2007; **23**(2): 471-489.

[16] Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, and Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society for Earthquake Engineering* 2009; **42**(2): 96-110.

[17] Bazzurro P. Probabilistic seismic demand analysis. Stanford University, Stanford, CA, 1998. 329. http://www.stanford.edu/group/rms/

[18] Aslani H and Miranda E. Probability-based Seismic Response Analysis. *Engineering Structures* 2005; **27**(8): 1151-1163.

[19] Aslani H and Miranda E. Fragility assessment of slab-column connections in existing non-ductile reinforced concrete buildings. *Journal of Earthquake Engineering* 2005; **9**(6): 777–804.

[20] Genz A. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing* 2004; **14**(3): 251-260.

[21] Drezner Z and Wesolowsky GO. On the computation of the bivariate normal inegral. *Journal of Statistical and Computational Simulation* 1990; **35**: 101-107.

[22] Chopra AK and Goel R. A modal pushover analysis procedure to estimate seismic demands for unsymmetric-plan buildings. *Earthquake Engineering and Structural Dynamics* 2004; **33**: 903–927.

[23] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: ground motion and structural response. *Earthquake Engineering and Structural Dynamics* 2008: (submitted).

[24] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: from structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2008: (submitted).

[25] Fajfar P. Structural analysis in earthquake engineering – A breakthrough of simplified nonlinear methods, in *12th European Conference on Earthquake Engineering*, London, UK., 2002.

[26] Baker JW and Cornell CA. Uncertainty Specification and Propagation for Loss Estimation using FOSM Method. Peer Report No 2003/07, 2003. 100. http://peer.berkeley.edu/publications/peer_reports.html

[27] Iervolino I, Manfredi G, and Cosenza E. Ground motion duration effects on nonlinear seismic response. *Earthquake Engineering and Structural Dynamics* 2006; **35**(1): 21-38, DOI: 10.1002/eqe.529.

[28] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra* 2007; **23**(2): 357-392, DOI: 10.1193/1.2723158.

[29] Ditlevsen O. *Uncertainty Modeling: with Applications to Multidimensional Civil Engineering Systems*. McGraw-Hill International Book Co.: New York; London, 1981; 412.

[30] Bradley BA. Epistemic uncertainty in component fragility functions. *Earthquake Spectra* 2009: (in press).

[31] Porter K. Fragility of Hydraulic Elevators for Use in Performance-Based Earthquake Engineering. *Earthquake Spectra* 2007; **23**(2): 459-469.

[32] Touran A. Probabilistic cost estimating with subjective correlations. *Journal of Construction Engineering and Management* 1993; **119**(1): 58-71.

[33] Hudak D and Maxwell M. A macro approach to estimating correlated random variables in engineering production projects. *Construction Management and Economics* 2007; **25**: 883-892.

[34] Touran A and Wiser EP. Monte carlo technique with correlated random variables. *Journal of Construction Engineering and Management* 1992; **118**(2): 258-272.

[35] Kulkarni RB, Youngs RR, and Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. *8th World Conference on Earthquake Engineering,*, San Francisco, CA, 1984.

[36] Kutner M, Nachtsheim C, Neter J, and Li W. *Applied Linear Statistical Models*. Fifth Edition ed. McGraw-Hill/Irwin: New York, 2005; 1396.

[37] Bradley BA. Seismic hazard epistemic uncertainty in the San Francisco bay area and its role in performance-based assessment. *Earthquake Spectra* 2009: (in press).

[38] Dolsek M. Incremental dynamic analysis with consideration of modelling uncertainties. *Earthquake Engineering and Structural Dynamics* 2009, DOI: 10.1002/eqe.869.

[39] Kramer SL, Arduino P, and Shin H. Using OpenSees for performance-based evaluation of bridges on liquefiable soils. Pacific Earthquake Engineering Research Centre, 2008. 198.

[40] Krawinkler H. Van Nuys hotel building testbed report: exercising seismic performance assessment. PEER 2005/11 Pacific Earthquake Engineering Research Centre, 2005. 264.

[41] Bradley BA. User manual for SLAT: Seismic Loss Assessment Tool version 1.14. *University of Canterbury Research Report No.2009-01*,Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand, 2009. 94.

[42] Bradley BA, Lee DS, Broughton R, and Price C. Efficient Evaluation of Performance-based Earthquake Engineering Equations. *Structural Safety* 2009; **31**(1): 65-74.

10.Seismic Hazard Epistemic Uncertainty in the San Francisco Bay Area and its Role in Performance-Based Assessment

Bradley BA. Seismic hazard epistemic uncertainty in the San Francisco bay area and its role in performance-based assessment. *Earthquake Spectra* 2009. (in press).

10.1 Abstract

This chapter investigates epistemic uncertainty in the results of seismic hazard analyses for the San Francisco bay area and their role in the broader picture of seismic performance assessment. Using the 2002 Working Group on California Earthquake Probabilities earthquake rupture forecast, epistemic uncertainty in the seismic hazard for several different intensity measures and sites in the San Francisco bay area is investigated. Normalization of the epistemic uncertainty for various sites and intensity measures illustrates that the uncertainty magnitude can be approximately estimated as a function of the mean exceedance probability. The distribution of the epistemic uncertainty is found to be dependent on the set of alternative ground motion prediction equations used, but is frequently well approximated by the lognormal distribution. The correlation in the hazard uncertainty is observed to be a function of the separation between the two different intensity levels, and a simple predictive equation is proposed based on the data analysed. Three methods for the propagation of seismic hazard epistemic uncertainty are compared and contrasted using an example of the 30-year collapse probability of a structure. It is observed that, for this example, epistemic uncertainty in the collapse capacity is more influential than that in the seismic hazard.

10.2 Introduction

Evaluation of the seismic risk of structures and facilities is burdened by significant uncertainties. In general such uncertainties can be classed as either aleatory or epistemic. Aleatory uncertainty is due to randomness, while epistemic uncertainty is due to the lack of knowledge of the process being observed. While the separation of uncertainties as being either aleatory or epistemic is not a trivial task, from a pragmatic viewpoint, the distinction should be made based on uncertainties which the analyst can and cannot reduce [1].

The distinctly different nature of these uncertainties implies that different methods should be employed to propagate them in uncertainty analyses. For probabilistic seismic hazard analysis (PSHA), in particular, where the goal is to determine the probability of exceeding some level of seismic intensity measure, aleatory uncertainties are considered explicitly in the computation of the seismic hazard and result in a single seismic hazard curve, while epistemic uncertainties lead to multiple hazard curves. Furthermore, aleatory uncertainties are, in theory, random such that given a long enough period of time all possible different values of the process will be realised. Epistemic uncertainties, on the other hand, represent different possibilities of a process which has only one true but currently unknown value and are commonly therefore non-ergodic [1, 2].

Treatment of seismic hazard and seismic risk using frameworks such as the Pacific Earthquake Engineering Research (PEER) centre performance-based earthquake engineering (PBEE) equation [3] can explicitly consider epistemic uncertainties in the seismic hazard at the site of a specific structure and propagate such uncertainty to other measures of seismic performance [e.g. 4].

This chapter investigates epistemic uncertainty in the results of seismic hazard analyses of the San Francisco bay area, and their role in the seismic performance assessment of structures and facilities. Using the 2002 Working Group on California Earthquake Probabilities (WGCEP02) earthquake rupture forecast (ERF) for the San Francisco bay area, the epistemic uncertainty in the seismic hazard for several different intensity measures at several sites is investigated. Based on the observed results some features of the epistemic uncertainty are characterised including: (i) uncertainty magnitude and variation with probability of exceedance; (ii) distribution of the probability of exceedance for a given intensity; (iii) correlation of the epistemic uncertainty in the exceedance probability at different intensity levels. Three methods of various complexity and input requirements for the

propagation of seismic hazard epistemic uncertainty to other seismic performance measures are compared and contrasted using an example of the collapse risk of a structure.

10.3 Consideration of epistemic uncertainties in PBEE

There exist many uncertainties in seismic hazard analysis, and while an exhaustive list is beyond the scope of this work (details can be found in, for example Kramer [5] and McGuire [6]), it is noted that such uncertainties can be classified as either relating to the prediction of earthquake ruptures, or the characterisation of the resulting ground motions. Typical epistemic uncertainties in earthquake prediction include [7]: (1) time-dependent nature of characteristic ruptures; (2) magnitude-area scaling relations; (3) fault segmentation endpoints; (4) seismogenic thickness; (5) fault slip rates; (6) relative frequency of various multi-segment ruptures; (7) amount of aseismic slip; (8) magnitude-frequency distributions; and (9) off-fault seismicity, among others. Given that an earthquake rupture occurs, epistemic uncertainties relating to the earthquake-induced ground motions observed at or near the ground surface may include: (1) Non-uniformity and occurrence time of slip across the rupture surface; (2) direction of fault rupture; (3) variation in ground motion attenuation with distance; (4) effects of geologic structures, such as basin and other topographic effects; and (5) effects of surficial soils, among others.

As will be seen, the magnitude of the aforementioned uncertainties is significant, and therefore it is prudent that they are considered via the use of logic trees [8], which is a discrete approach used in contemporary PSHA for handling epistemic uncertainties.

Remembering that seismic hazard analyses are used as an input to determine the seismic risk of structures and facilities, then epistemic uncertainties in seismic hazard will result in uncertainty in the value of seismic performance measures which are dependent on the seismic hazard. The following section investigates the characteristics of epistemic uncertainties for a variety of ground motion intensity measures (IM's) and sites in the San Francisco bay area; while the last section investigates three methods by which epistemic uncertainties in seismic hazard can be propagated to seismic risk measures.

10.4 Characterisation of epistemic uncertainty

Characterisation of epistemic uncertainties in seismic hazard requires determination of the magnitude, distribution and correlation structure of the uncertainty. While inevitably each

of these characteristics will be specific to a particular site, several important details can be obtained by considering several sites and ground motion IM's in regions where the epistemic uncertainties are relatively well researched. Herein use is made of the detailed ERF developed for the San Francisco bay area [7].

10.4.1 The WGCEP02 earthquake rupture forecast (ERF)

The Working Group on California Earthquake Probabilities [7], or WGCEP02 herein, developed a time-dependent ERF for the San Francisco bay area which is arguably the most sophisticated ERF ever developed [9]. This ERF contains all epistemic uncertainties related to earthquake occurrence discussed in the previous section. The sophisticated nature of the ERF meant that it was not feasible to quantify the epistemic uncertainty by directly considering all of the end nodes of the logic tree, but to instead use a Monte Carlo procedure based on the relative weights of the various branches [7]. The WGCEP02 ERF has been implemented in the OpenSHA framework [10], an open-source code for seismic hazard analysis and was used in this study. A typical seismic hazard curve for PGA in San Francisco is illustrated in Figure 10-1. It is important that such a rigorous example has been considered in this study as will be later discussed.

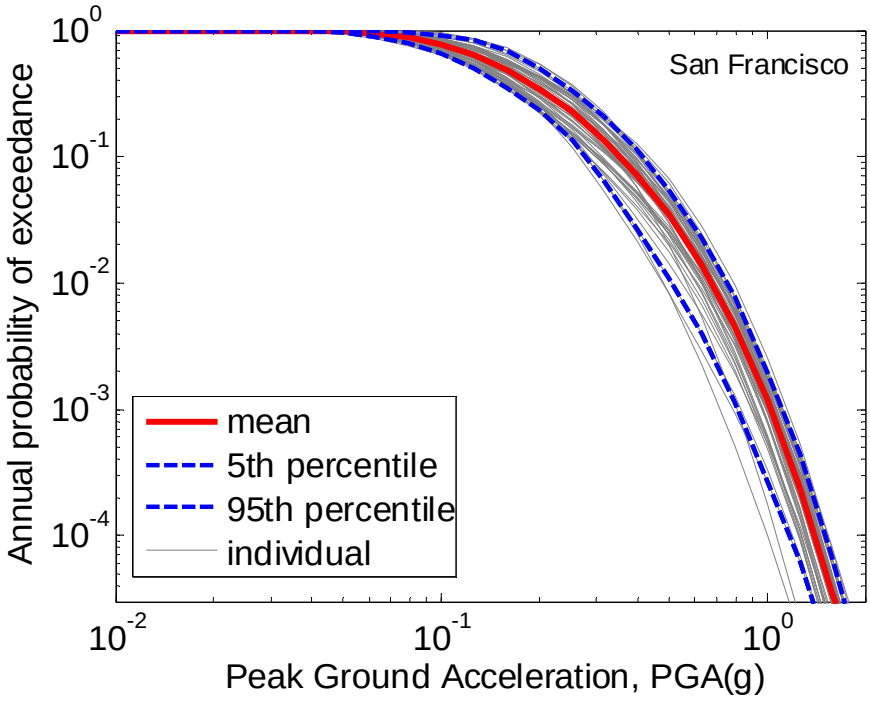


Figure 10-1: Illustration of earthquake rupture forecast (ERF) epistemic uncertainty in the peak ground acceleration hazard curve for a site ($V_{s(30)}$=760 m/s) in the San Francisco bay area using the Campbell and Bozorgina [11] prediction equation.

10.4.2 Geographical locations considered

Four different sites were investigated in order to consider different levels of seismicity and different dominant faults. The sites were San Francisco (37.80°N,122.42°W); Stockton (37.90°N,121.25°W); Sacramento (38.52°N,121.50°W); and San Jose (37.37°N,121.93°W). The mean seismic hazard curves for the different sites are shown in Figure 10-2 (all sites located on soil with $V_{s(30)}$=760m/s). These four sites span a wide range of seismic hazard with 6.1% and 1.2% in 30 year exceedance values in the region of 0.1g and 0.2g in Stockton and Sacramento, and 0.5g and 0.8g in San Francisco and San Jose (although the ERF is strictly time-dependent, for the purpose of comparison these two probabilities of exceedance are approximately equivalent to 10% and 2% in 50 years using the Poisson assumption). A time span of 30 years is used herein as was adopted in the WGCEP02 study. While two of the regions each represent moderate-to-low seismicity and high seismicity, different faults will dominate the hazard at each site due to their different geographical locations.

Because interest in this study relates to epistemic uncertainties, it is noted that the seismic hazard curves presented herein are computed neglecting background seismicity which has low epistemic uncertainty (it is highly constrained by instrumental seismicity). The Gutenberg-Richter 'tail' [7] associated with the magnitude distributions of individual faults is however considered.

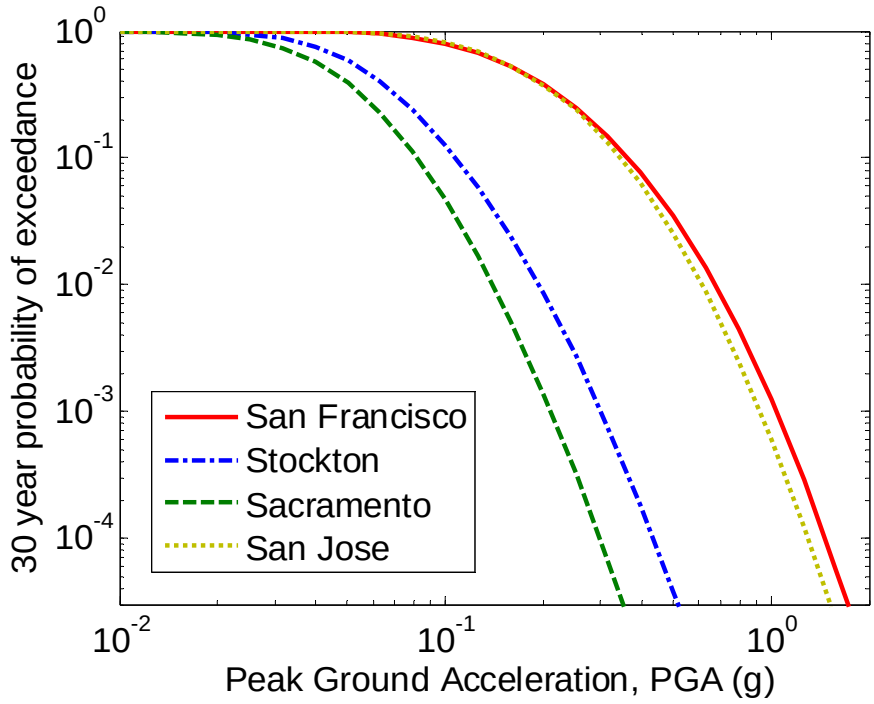


Figure 10-2: Mean hazard curves of the four different sites considered.

10.4.3 Different ground motion prediction equations considered

The WGCEP02 ERF, as the name implies, provides only the spatial and temporal occurrence of earthquake ruptures and not the resulting ground motions at the four considered sites. For this purpose ground motion prediction equations (GMPE's) are required. In this study two sets of GMPE's were used. The first set of GMPE's were those used in the 2002 update of the United States National Seismic Hazard Maps [12], specifically: Campbell and Bozorgina [11], Boore et al. [13], Abrahamson and Silva [14], and Sadigh et al. [15]. For brevity these models are referred to as CB03, BJF97, AS97, and S97, respectively. The second set of GMPE's were developed as part of the Next Generation Attenuation (NGA) project which included: Campbell and Bozorgina [16], Boore and Atkinson [17], and Abrahamson and Silva [18]. This second set is referred to as CB08, BA08 and AS08, respectively. It is noted that the other two empirical GMPE's developed in the same phase of the NGA project [19, 20] were not currently available in OpenSHA at the time this study was conducted.

Figure 10-3 illustrates the mean seismic hazard curves obtained for one-second spectral acceleration in San Francisco using the two different sets of GMPE's and the WGCEP02 ERF. Clearly in the case of Figure 10-3a, the magnitude of epistemic uncertainty due to

different GMPE's will be of the same order of magnitude as the epistemic uncertainty due to the prediction of earthquake occurrence (e.g. Figure 10-1). In the case of Figure 10-3b however, the difference between the hazard curves obtained using the CB08, BA08 and AS08 models is significantly less (although two of the NGA models were not considered, comparisons of all of the models indicate that the same conclusion would be drawn [21]). In the results to follow, each of the prediction equations in the two sets of GMPE's were given equal weighting (i.e. 25% for each in the first, and 33% for each in the second).

Four different response spectral quantities, namely: PGA, S_a(0.5s), S_a(1.0s), and S_a(2.0s) (all 5% damped) were considered to investigate the variation of epistemic uncertainties with ground motion intensity measure type.

10.4.4 Magnitude of epistemic uncertainty

The 'magnitude' of epistemic uncertainty, as referred to herein, represents the size of the epistemic uncertainty in the ground motion hazard curves. The magnitude of epistemic uncertainty relates to the level of scientific uncertainty in the prediction of the seismic hazard, and is comprised of uncertainty in the occurrence of earthquake ruptures, and uncertainty in their resulting ground motions (i.e. uncertainty in the ERF and ground motion prediction, GMP). Herein epistemic uncertainties in the ERF were obtained by using 50 Monte Carlo simulations (for each GMPE) of the WGCEP02 logic tree, while uncertainty in GMP was considered via the use of multiple GMPE's, each with equal weighting. The number of Monte Carlo simulations was based on a compromise between adequate representation of the epistemic uncertainty (i.e. convergence in the magnitude of the uncertainty) and computational time.

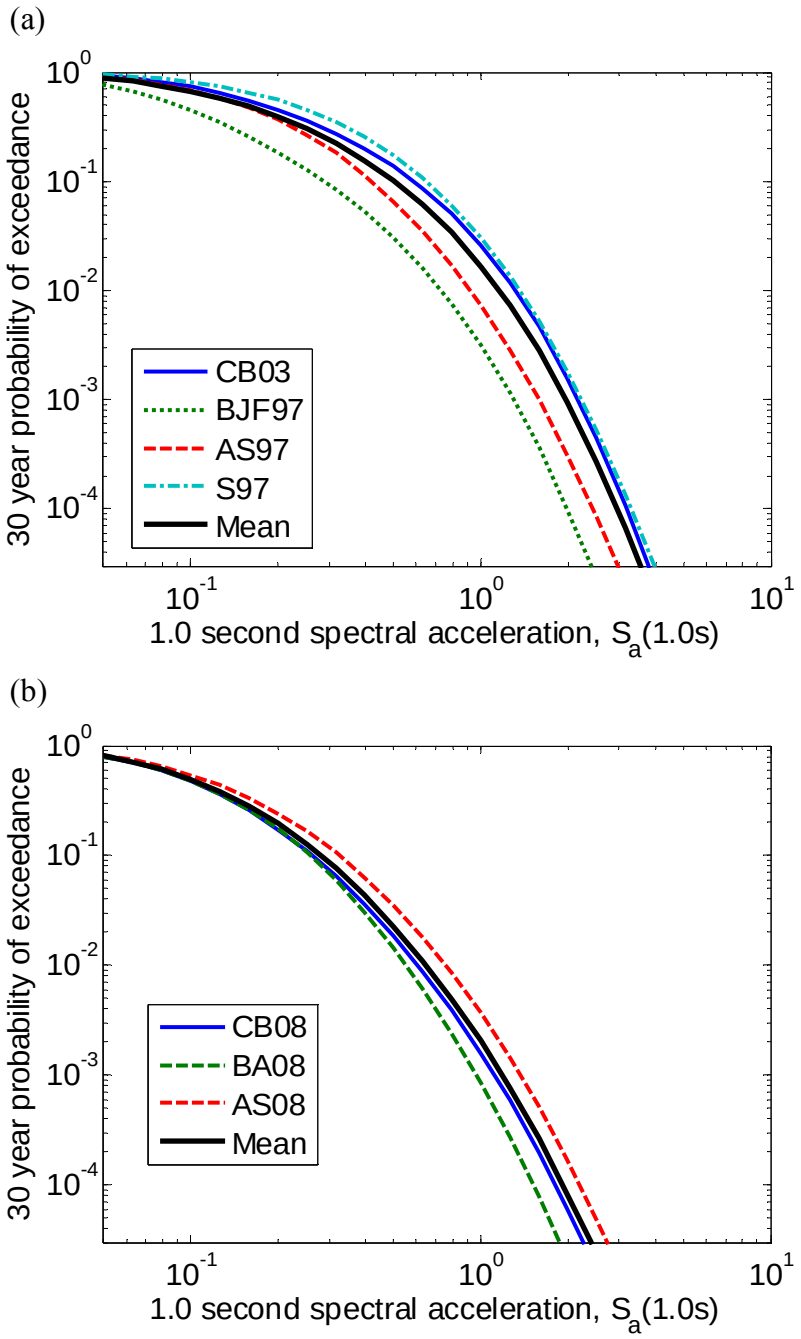


Figure 10-3: Illustration of the effect of various attenuation relations on the mean hazard curve in San Francisco using the two different ground motion prediction equation sets.

As the hazard curves for different sites, ground motion prediction equations, and ground motion intensity will all be unique then to study the trends in the epistemic uncertainty it was necessary to normalise the data so it can be presented together. For each seismic hazard curve the epistemic uncertainty was quantified using the lognormal standard deviation in the probability of exceedance, $\sigma_{\ln P(IM>im)}$, as a function of IM. The lognormal standard deviation was used because as the following section illustrates the lognormal distribution is a good approximation for the epistemic uncertainty in the hazard for the majority of the sites and

intensity measures considered. For each IM value at which $\sigma_{\ln P(IM>im)}$ was computed, the mean exceedance probability, $\mu_{\ln P(IM>im)}$, was also computed thus yielding pairs of $\left(\mu_{\ln P(IM>im)}, \sigma_{\ln P(IM>im)}\right)$ data for each site and intensity measure. Figure 10-4a and Figure 10-4b illustrate the variation in $\sigma_{\ln P(IM>im)}$ with $\mu_{\ln P(IM>im)}$ for the four different sites and ground motion IMs (i.e. there are 4 sites x 4 intensity measures = 16 lines for each GMPE in Figure 10-4a and Figure 10-4b) when only ERF uncertainty is considered. While there is obviously scatter in the results for the different sites and IM's, it can be seen that the normalisation allows, in part, for the different seismicity and ground motion measures, and that there is a clear trend for increasing epistemic uncertainty as the exceedance probability reduces. Once the data were plotted in the form shown in Figure 10-4a and Figure 10-4b, trends in the magnitude of epistemic uncertainty for the different sites, ground motion measures, and GMPE's were investigated. Figure 10-4a and Figure 10-4b illustrate the deaggregation of the data based on the different GMPE's. In the case of Figure 10-4a there is a clear trend that ERF uncertainty gives larger uncertainty in the seismic hazard when the S97 and AS97 models are used compared to the BJF97 and CB03 models. Figure 10-4b does not indicate any dependence of epistemic uncertainty magnitude on the CB08, BA08, or AS08 models, for all of which the magnitude is similar to the BJF97 and CB03 models. No obvious dependence on geographical site or intensity measure type was observed when only a single GMPE was used.

Figure 10-4c and Figure 10-4d illustrate the magnitude of epistemic uncertainty when considering uncertainty in the ERF and GMP using various GMPE's for the four sites and four ground motion IM types. Similar, to Figure 10-4a and Figure 10-4b the magnitude of the epistemic uncertainty increases with reducing probability of exceedance. Also, the magnitude of epistemic uncertainty when considering uncertainty in both ERF and GMP is significantly larger than when considering ERF uncertainty alone. Table 10-1 illustrates at the 1.2% in 30 year probability of exceedance, the range of values for the dispersion in the seismic hazard based on the results in Figure 10-4, as well as the range of the ratio of the 84th percentile to median exceedance probability for this given mean exceedance probability. As can be seen from Figure 10-4a and Figure 10-4b and the second column of Table 10-1, the dispersion with only ERF uncertainty is similar when using the two different sets of ground motion prediction equations. However, when considering GMP uncertainty, there is a significant increase in the magnitude of epistemic uncertainty using the first set of GMPE's compared to the second. In particular, the deaggregation of lines based on ground motion IM type in Figure 10-4c reveals

that the uncertainty is much lower when predicting PGA than the other three spectral quantities (while there is also evidence in Figure 10-4d of lower uncertainty in PGA, it is less pronounced).

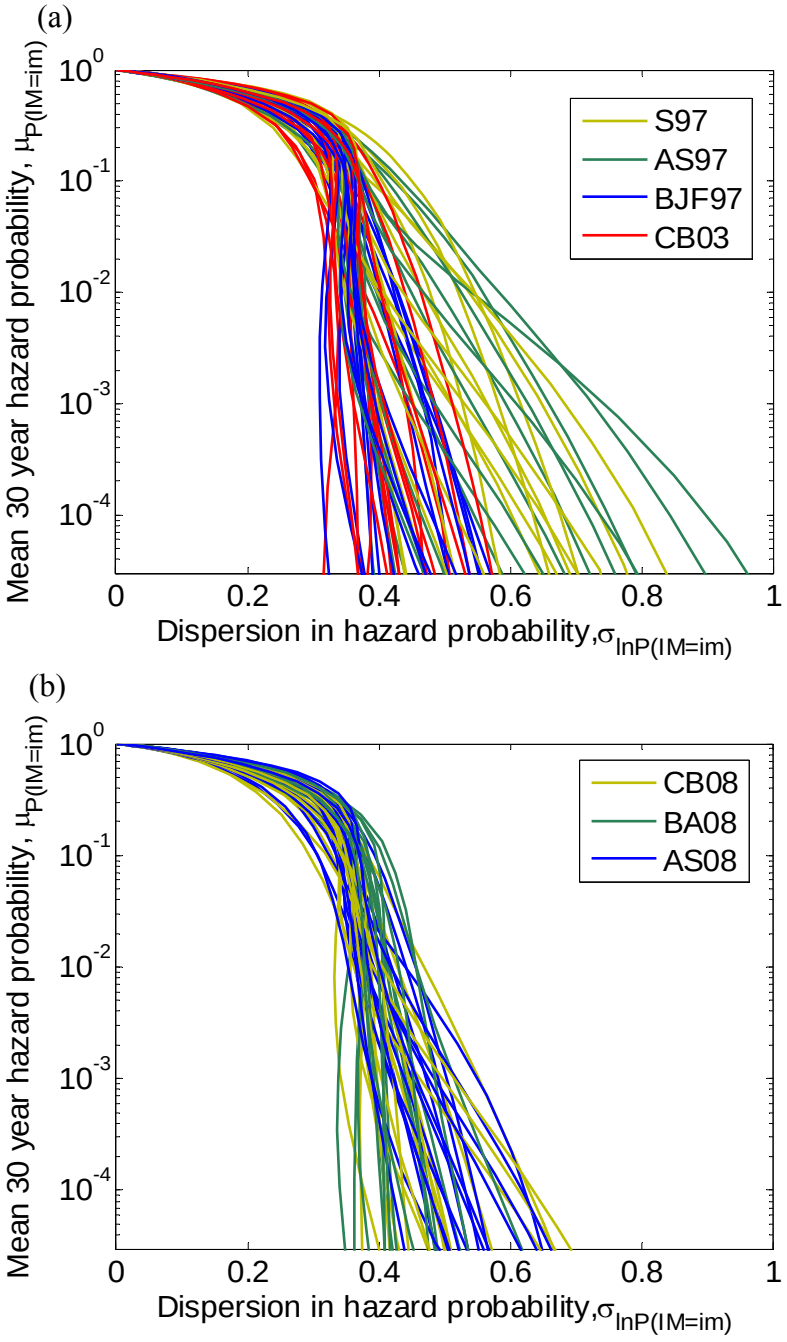


Figure 10-4: Magnitude of epistemic uncertainty in ground motion hazard estimates using: (a)&(b) only single ground motion prediction equation; and (c)&(d) using multiple ground motion prediction equations.

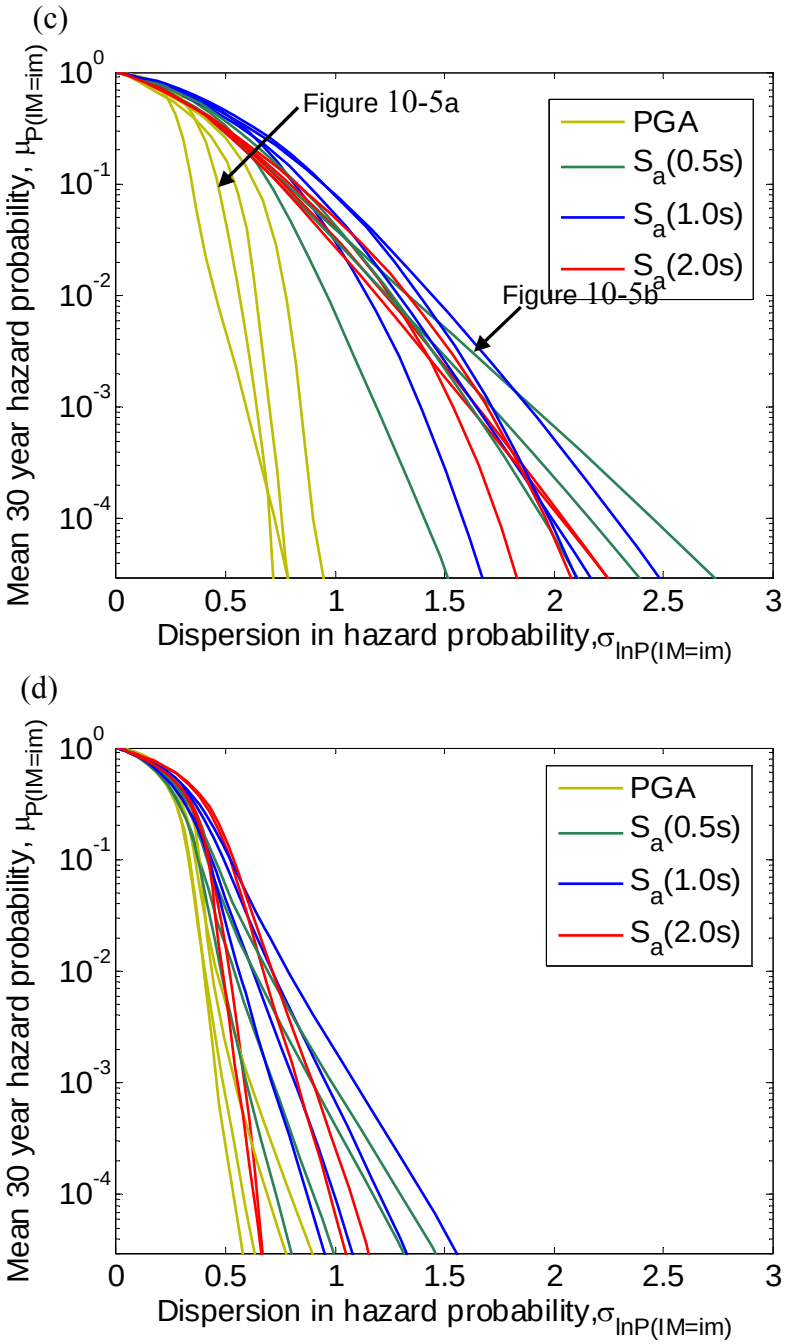


Figure 10-4 continued.

Table 10-1: Dispersion of epistemic uncertainty at 1.2% in 30 years probability of exceedance

Ground motion equations	Epistemic uncertainty in ERF only	Epistemic uncertainty in ERF and GMP
CB03, BJF97, AS97, S97	0.3-0.6 (1.35-1.82)*	0.5-1.5 (1.65-4.50)
CB08, BA08, AS08	0.35-0.5 (1.42-1.65)	0.5-0.8 (1.65-2.20)

* values in brackets give the ratio of the 84th percentile to median hazard (=exp(σ))

Figure 10-5 illustrates the seismic hazard curves for two different sites and ground motion measures which are annotated in Figure 10-4c. It can be seen in Figure 10-5a that most of the seismic hazard curves from the different GMPE's are overlapping, an indication that the differences in the ground motion prediction equations are of a similar order as the uncertainty in the ERF itself. Figure 10-5b illustrates the case where there is significantly larger disparity in the seismic hazard curves using the different GMPE's. In particular the BJF97 model gives a significantly lower hazard than the other three models. The second set of GMPE's (i.e. CB08, BA08, AS08) give a smaller difference in seismic hazard (as evident from Figure 10-3), and therefore the total epistemic uncertainty using these equations is less than using the first set of GMPE's. Campbell and Bozorgina [22] however suggest that because of the similarity in the predictions of the NGA models due to similar theories and empirical data, additional models for epistemic uncertainty in GMP could be used to better represent the true epistemic uncertainty. Hence, the values in Figure 10-4d can be considered lower than the true epistemic uncertainty.

As noted by Abrahamson [23] logic trees are commonly interpreted to represent the uncertainty in seismic hazard analysis, however in reality they represent the range of available alternative scientific models. A consequence of this is that using available models for a site with little or no data will indicate smaller epistemic uncertainty compared with a well studied site with many available models, when clearly the poorly studied site will have a larger epistemic uncertainty [23]. Hence the results shown in Figure 10-4, which represent the epistemic uncertainty for sites with comprehensive alternative scientific models, can be used as somewhat of a lower bound for such unstudied sites. Additionally, as current ground motion hazard maps for the San Francisco bay area provide only the mean exceedance rate or probability for a given ground motion intensity then Figure 10-4 also provides a means to approximately consider the magnitude of seismic hazard epistemic uncertainties in the performance-based assessment of facilities if site-specific PSHA details (i.e. logic-tree results) are not available.

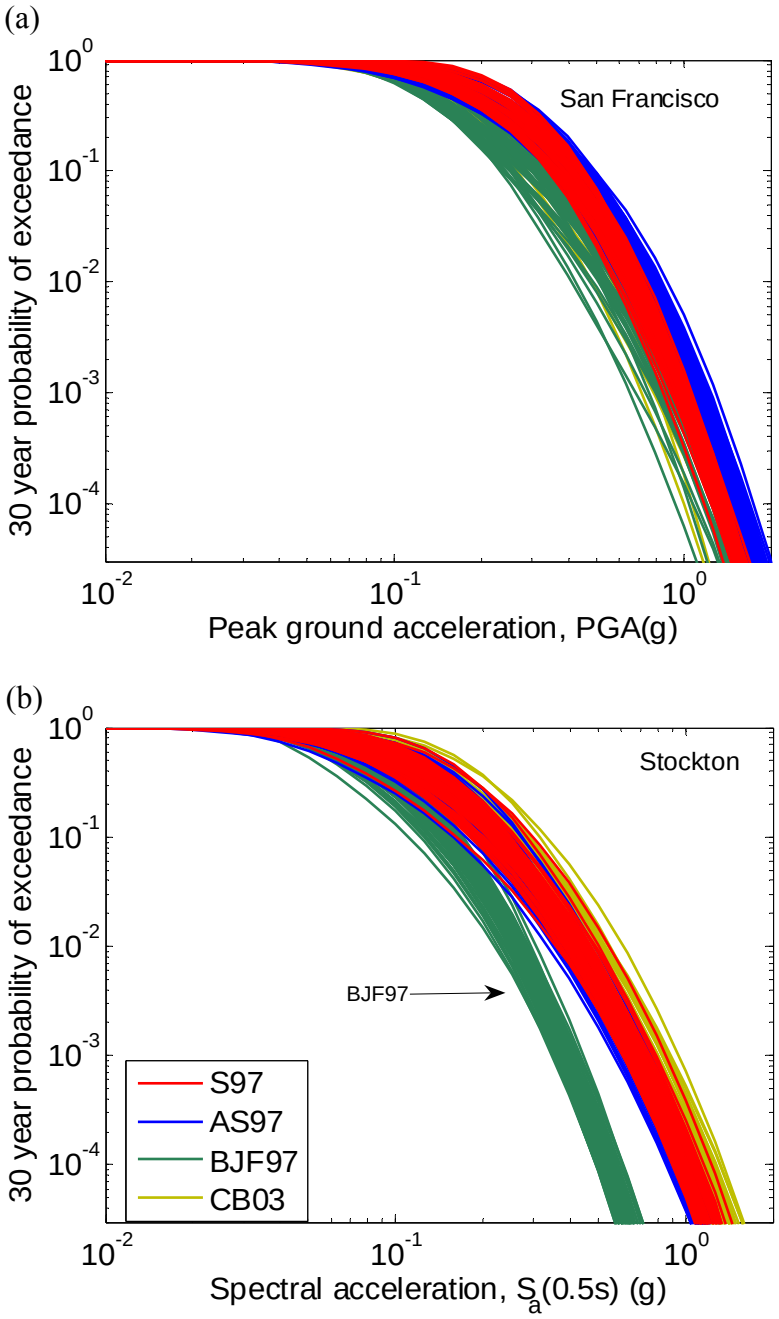


Figure 10-5: Example of ground motion hazard curves using different ground motion prediction equations: (a) where 'inconsistency' does not occur; and (b) where 'inconsistency' does occur.

10.4.5 Distribution of epistemic uncertainty

As noted earlier, each different PSHA performed via a single branch of a logic tree results in a single ground motion hazard curve which contains only aleatory uncertainty. All of the different possible combinations of the logic tree give different possible hazard curves, which represent the epistemic uncertainty. The mean hazard curve can be determined as the

weighted average of the probability of exceedance values for a given IM from each of the different hazard curves. The mean hazard curve is typically that presented [e.g. 12, 24] for the design of structures, although some note that other possible options should be considered [25-27]. For each level of ground motion intensity, IM, the *n* different ground motion hazard curves from the logic tree provide *n* different values for the probability (or rate depending on the PSHA formulation used [10]) of exceeding that level of IM. Thus from the *n* different probabilities it is possible to determine the empirical distribution of the exceedance probability of a specified IM value.

When investigating the empirical distributions for the epistemic uncertainty in the seismic hazard curves for a single ground motion prediction equation (i.e. only epistemic uncertainty in the ERF) it was found that over the range of different sites, intensity measures, and intensity measure levels the lognormal distribution frequently provided an acceptable fit based on the Kolmogorov-Smirnov goodness-of-fit test at the 95% confidence level [28]. This observation may be a result of the assumption that fault dimensions and slip rates in the ERF are typically defined as having normal or lognormal distributions [7].

Figure 10-6a illustrates the cumulative distribution for the epistemic uncertainty when predicting a PGA = 1.0g at the San Francisco site using the CB03, BFJ97, AS97, S97 models, while Figure 10-6c illustrates the probability values from the 50 Monte Carlo simulations conducted for each GMPE. While the difference in the mean probability value from each of the GMPE's range from $6x10^{-4}$ – $2x10^{-3}$, it is clear that there is significant overlap of the individual simulations from the different GMPE's, and that this results in a total epistemic uncertainty which is well approximated by the lognormal distribution (Figure 10-6a).

Figure 10-6b and Figure 10-6d show the same results as Figure 10-6a and Figure 10-6c, for the distribution of $S_a(0.5s)$=0.5g at the Stockton site. Clearly, the BJF97 model leads to significantly lower exceedance probability values than the other three models, with the deviation becoming more apparent as the probability of exceedance reduces, and resulting in a cumulative distribution which departs significantly from the lognormal distribution. Figure 10-5b illustrates that using the BJF97 model for $S_a(0.5s)$ leads to lower hazard estimates over a wide range of IM values, but Figure 10-5a illustrates that it is in agreement for other IM types. This may lead one to suggest that the logic tree weights applied to the different GMPE's should be a function of the IM type and value considered, as advocated by Scherbaum et al. [29]. Although not shown here, when using the CB08, BA08 and AS08 models the distribution of epistemic uncertainty resembled that for Figure 10-6a and Figure 10-6c, which is to be expected as it is apparent from Figure 10-3b that there is little difference

between the mean hazard estimates using this set of GMPE's.

The observation that the lognormal distribution provides a good representation of the epistemic uncertainty is desirable because the lognormal distribution is completely defined by its mean and standard deviation. The mean is that which is commonly provided in literature, while the standard deviation could be estimated based on the results of the previous section (if site-specific analysis is not viable). However, it should be kept in mind that the above observations apply to the specific data examined and may vary when other tectonic regimes and GMPE's are considered.

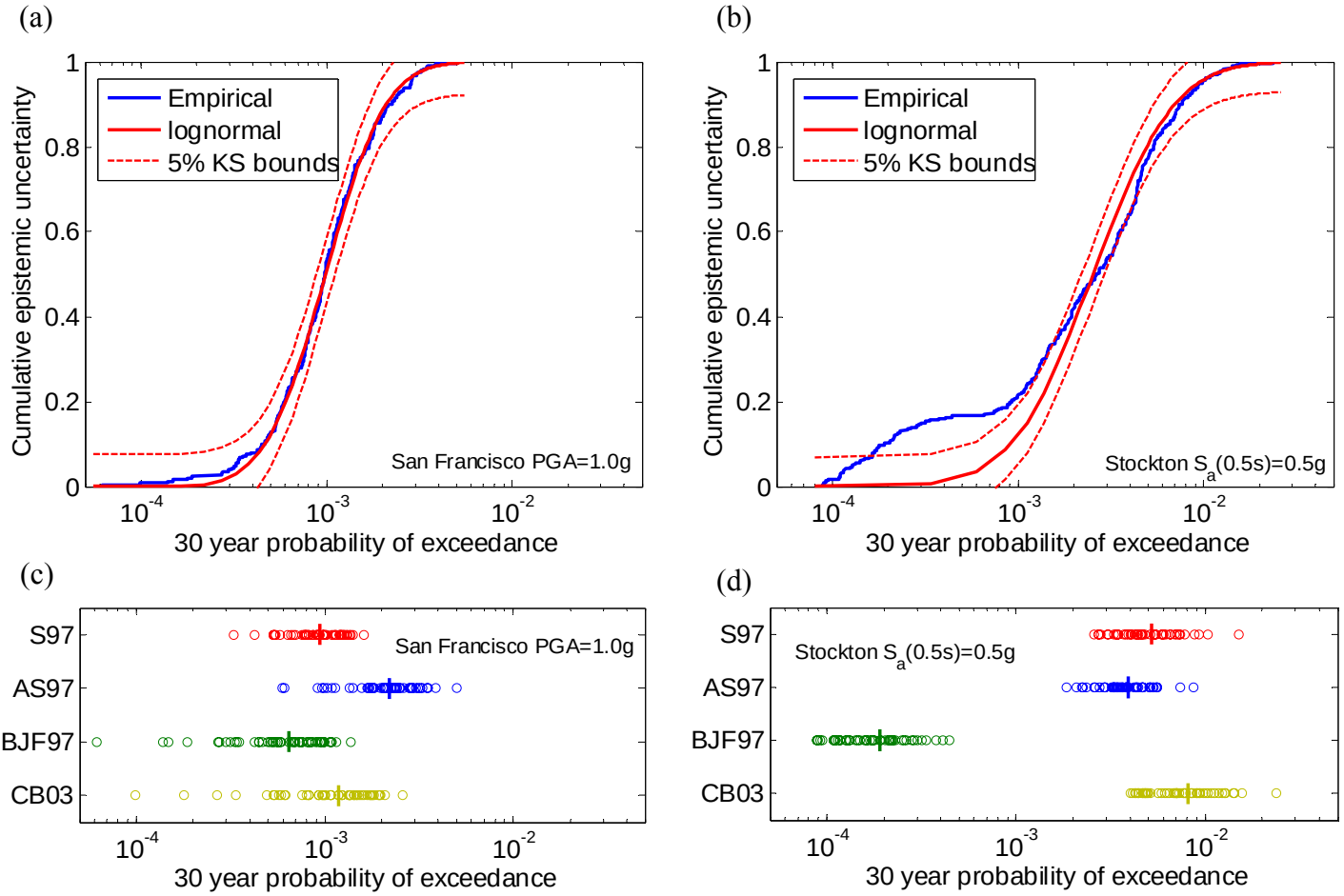


Figure 10-6: Distribution of epistemic uncertainty in seismic hazard curves: (a)&(b) Kolmogorov-Smirnov goodness-of-fit tests for the lognormal distribution; and (c)&(d) distribution of simulation data from different ground motion prediction equations.

10.4.6 Correlation

The correlation within the epistemic uncertainty at various levels of ground motion intensity, as will be seen, is important in the process of propagation of the uncertainty in the

performance-based seismic risk assessment of structures and facilities. The correlation of epistemic uncertainties in the ground motion hazard relates to the dependence of the epistemic uncertainty at different levels of IM. Figure 10-7 illustrates the simulation of ground motion hazard curves using the lower and upper bound correlations of zero and one, respectively. Comparison with the hazard curves presented in the remainder of this chapter (e.g. Figure 10-1) illustrates that the correlation structure of ground motion hazard curves observed in practice is much closer to the perfect correlation case than being uncorrelated. In fact, Figure 10-7a illustrates that if care is not taken regarding the correlation coefficient it is possible to generate hazard curves which are not one-to-one (i.e. not monotonically decreasing for increasing intensity) and therefore impossible.

For a single site and ground motion IM it is possible to estimate the correlation between probabilities of exceedance at two different IM levels using the sample correlation coefficient:

$$\rho_{i,j} = \frac{\sum_{k=1}^{n}\left(\ln P(im_i)_k - \overline{\ln P(im_i)}\right)\left(\ln P(im_j)_k - \overline{\ln P(im_j)}\right)}{\sqrt{\sum_{k=1}^{n}\left(\ln P(im_i)_k - \overline{\ln P(im_i)}\right)^2}\sqrt{\sum_{k=1}^{n}\left(\ln P(im_j)_k - \overline{\ln P(im_j)}\right)^2}} = \sum_{k=1}^{n}\varepsilon_{i,k}\varepsilon_{j,k} \qquad (10\text{-}1)$$

where $\rho_{i,j}$ is the correlation between $\ln P(IM > im_i)$ and $\ln P(IM > im_j)$; $P(im_i)_k$ is the k[th] value of $P(IM > im_i)$; $\overline{\ln P(im_i)}$ is the sample mean of $P(im_i)$; and $\varepsilon_{i,k} = \left(\ln P(im_i)_k - \overline{\ln P(im_i)}\right) \Big/ \sqrt{\sum_{k=1}^{n}\left(\ln P(im_i)_k - \overline{\ln P(im_i)}\right)^2}$ is the 'residual' of $P(im_i)_k$. The correlation between the logarithms was used because the distribution of $P(IM > im_i)$ was shown to be approximately lognormal in the previous section.

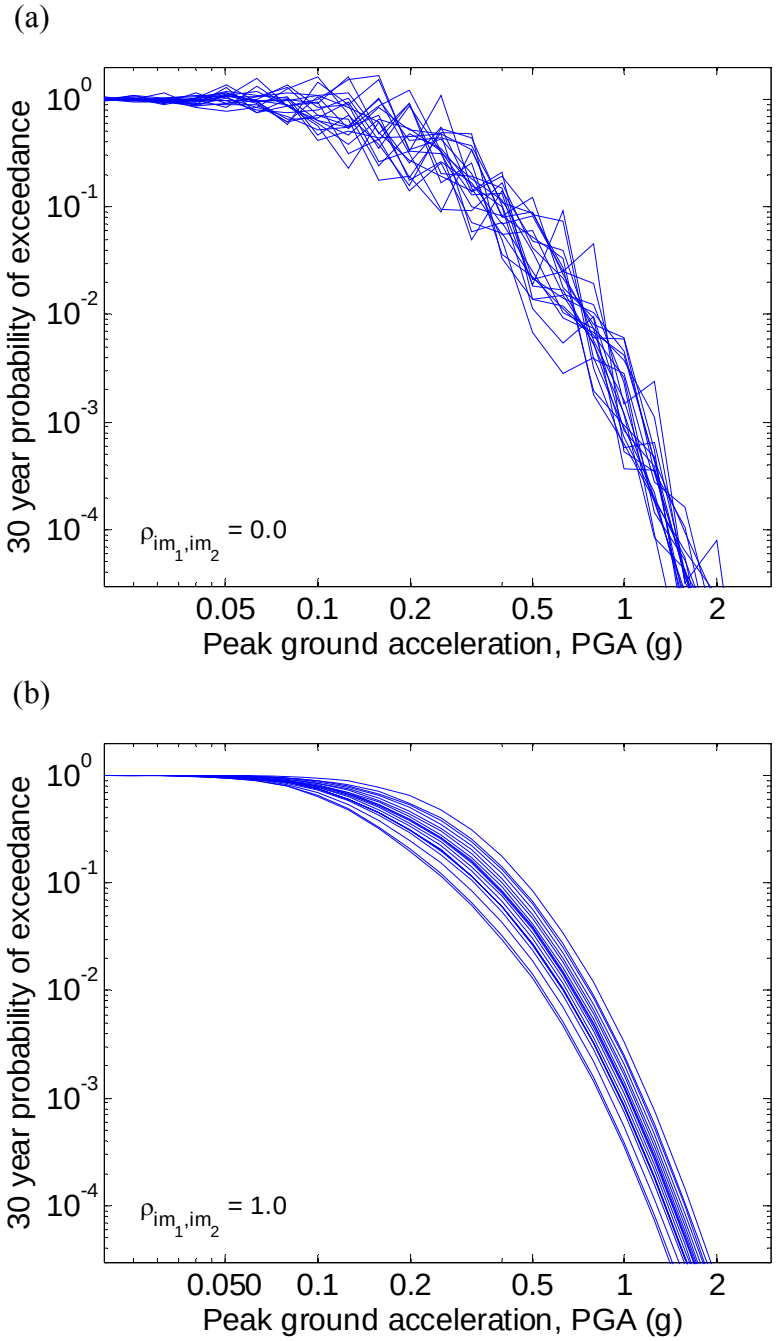


Figure 10-7: Effect of correlation on ground motion hazard generation.

Figure 10-8 illustrates two typical plots of the residuals using the hazard data for PGA at the San Francisco site. Figure 10-8a illustrates that for similar IM values there exists a high correlation, while Figure 10-8b illustrates for IM values which are significantly different the correlation is weak. This seems relatively intuitive that the dependence decreases as the separation between im_i and im_j increases. Figure 10-9a illustrates this trend, where for three different values of im_i the variation in the correlation is computed at multiple im_j values. A simple predictive equation for the correlation coefficient was obtained by transforming the data from the form shown in Figure 10-9a to one where the abscissa is normalised by the

value of im_i. Figure 10-9b illustrates the transformed San Francisco PGA data, the mean obtained using non-parametric regression, and the parametric equation given by:

$$\rho_{i,j} = 1.0 - 0.12\left[\ln\left(im_i \diagup im_j\right)\right]^2 \qquad 0.25 \le \frac{im_i}{im_j} \le 4.0 \tag{10-2}$$

Comparisons of the empirical form of the correlation coefficient defined by Equation (10-2) with data from the four different sites and four IMs was found to be similar to that shown in Figure 10-9b. As mentioned previously, it is possible that Monte Carlo simulation of seismic hazard curves may produce physically impossible results (e.g. Figure 10-7a). While this is still a theoretical possibility using Equation (10-2), because the correlation coefficient is approximately 1.0 for similar IM values it is extremely unlikely. For example, if $im_i = 0.1$ and $im_j = 0.2$, Equation (10-2) gives $\rho_{i,j} = 0.94$, and Figure 10-3 illustrates that the mean exceedance probabilities for these IM values are ~0.5 and ~0.05, respectively (i.e. a factor of 10 different). Using Monte Carlo simulation for this scenario gives a probability of less than 1×10^{-8} that $P(im_i) < P(im_j)$. Thus Monte Carlo simulation of seismic hazard curves is not a practical problem when correlations are appropriately considered.

It should be noted that when performing performance-based computations which require this correlation (discussed in the following section) the major contribution to the value of the performance measure integral occurs over a 'small' region of the integration variable (e.g. Bradley and Dhakal [30, Figure 4]) such that typically only the correlation over the range im_i/im_j = 0.3-3.0 will be important. Because of the high correlation over this small range of the integration variable the subsequent section illustrates that there is little difference between using the correlation model of Equation (10-2) and the perfect correlation assumption.

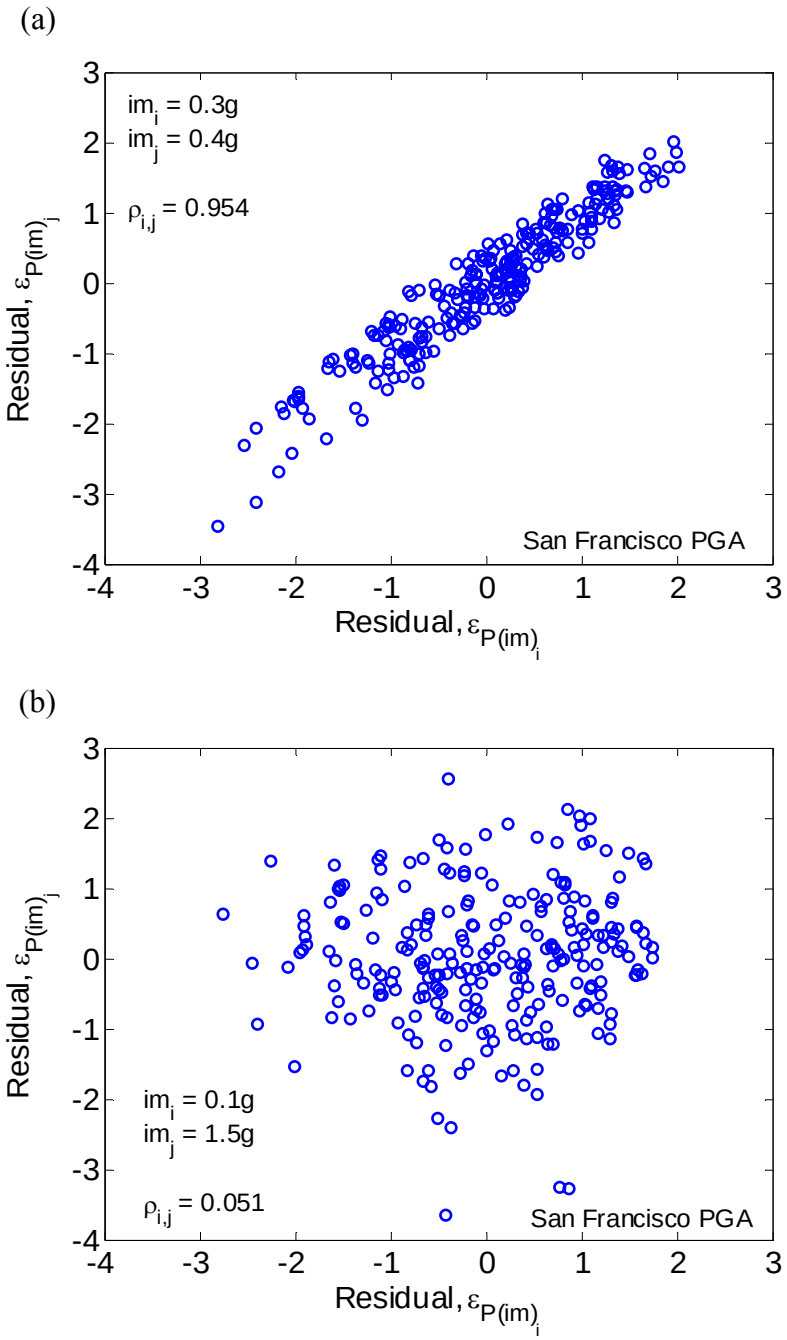


Figure 10-8: Illustration of correlation of epistemic uncertainty in the ground motion hazard curve for: (a) intensity measure values close in absolute magnitude; and (b) intensity measure values distant in absolute magnitude.

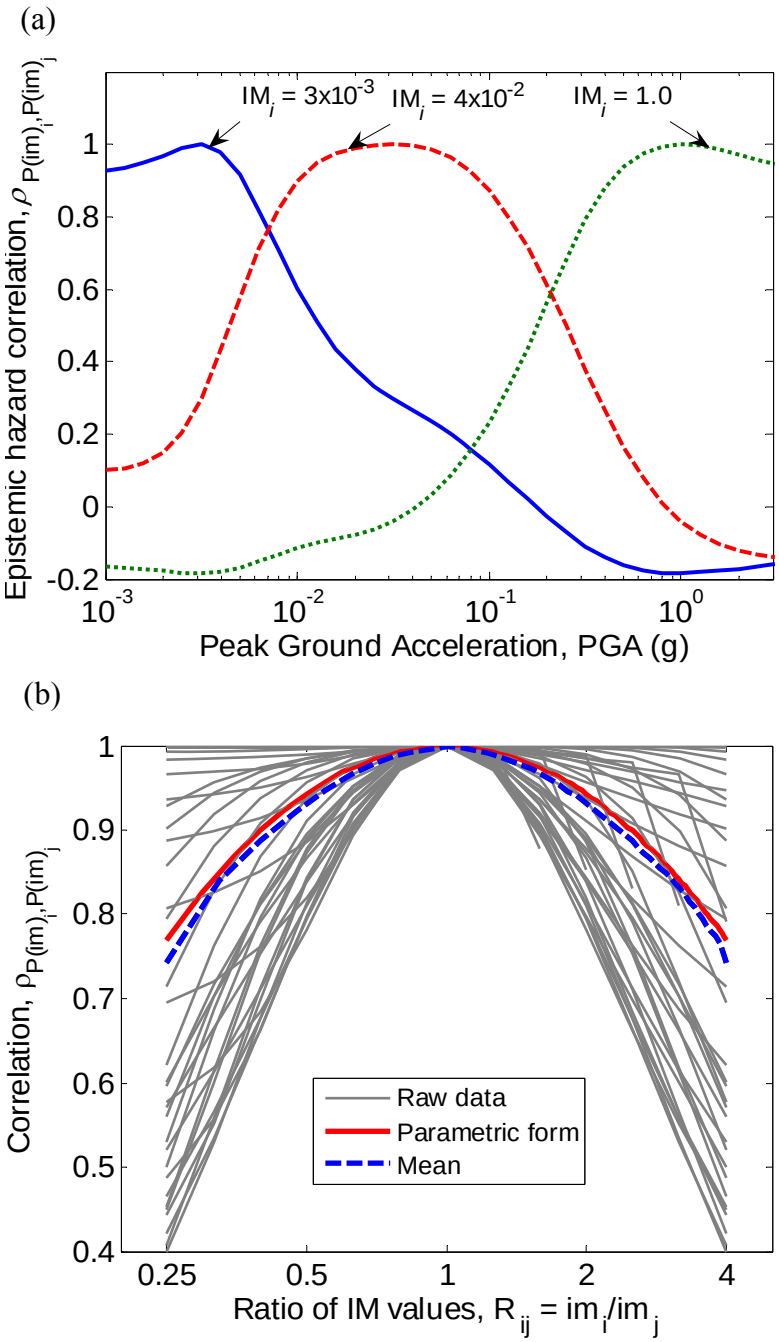


Figure 10-9: Correlation of epistemic uncertainty between different intensity measure values: (a) for three different intensity measure values; and (b) for all intensity measure values after normalization.

10.5 Propagation of epistemic uncertainty

In practice, determination of the ground motion hazard at a site is the first step in the seismic design or assessment of some engineered facility. Typically, the ground motion hazard is used to determine the level of seismic intensity a structure will be subject to for a given probability of exceedance. Some form of seismic response analysis is then performed

to determine the response of the structure to this level of ground motion and consequences (repair cost, injuries, business disruption) associated with the seismic response are explicitly or implicitly considered. The PEER PBEE framework provides a robust methodology for quantification of the seismic performance of structures, utilizing the theorem of total probability in the same fashion as that to compute seismic hazard within PSHA.

For simplicity, attention here will be given to computation of the probability of collapse in some time interval, which is given by (in continuous and discrete forms):

$$P_C = \int_{IM} P\left(C \middle| IM = im\right) \left| \frac{dP\left(IM > im\right)}{dIM} \right| dIM = \sum_{i=1}^{n} w_i P\left(C \middle| IM = im_i\right) P\left(IM = im_i\right) \tag{10-3}$$

where $P\left(C \middle| IM = im\right)$ is the probability of collapse given $IM = im$; $P\left(IM > im\right)$ is the ground motion hazard for the given time interval; and w_i is an integration weighting which will depend on the numerical integration procedure used.

Equation (10-3) illustrates that the collapse probability is obtained by combining the collapse fragility (obtained from seismic response analyses) and the ground motion hazard. In both of these relationships epistemic uncertainties exist, which should be propagated in Equation (10-3) to compute the uncertainty in the collapse probability. Below three methods for propagation of these uncertainties which cover a range of complexity and accuracy are discussed.

10.6 Parametric second moment method

Based on the discrete form of Equation (10-3) it is possible to determine the uncertainty in the 30-year collapse probability using the method of moments [4, 28]. Firstly, the mean (or best-estimate) of the collapse probability is given by the expectation of Equation (10-3):

$$E\left[P_C\right] = E\left[\sum_{i=1}^{n} w_i P\left(C \middle| IM = im_i\right) P\left(IM = im_i\right)\right] \tag{10-4}$$

Making use of the linearity of the expectation operator and noting that the epistemic uncertainty in $P\left(C \middle| IM = im_i\right)$ and $P\left(IM = im_i\right)$ is uncorrelated one obtains:

$$E\left[P_C\right] = \sum_{i=1}^{n} w_i \mu_{P\left(C \middle| IM = im_i\right)} \mu_{P\left(IM = im_i\right)} \tag{10-5}$$

where μ_Z is the mean of Z. As the discrete form of Equation (10-3) is of the form $\sum_{i=1}^{n} X_i Y_i$

(i.e. the summation of the product of uncorrelated random variables) then it can be shown (i.e. Appendix C) that the variance of the collapse probability is given by:

$$Var[P_C] = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \left\{ \sigma_{C_i,C_j}\sigma_{IM_i,IM_j} + \mu_{IM_i}\mu_{IM_j}\sigma_{C_i,C_j} + \mu_{C_i}\mu_{C_j}\sigma_{IM_i,IM_j} \right\} \quad (10\text{-}6)$$

where C_i and IM_i are shorthand notation for $P(C|IM = im_i)$ and $P(IM = im_i)$, respectively; and σ_{Z_i,Z_j} is the covariance between Z_i and Z_j.

Thus the second moment method makes it possible to compute the mean and variance in the 30-year collapse probability, with only knowledge of mean and covariance of the seismic hazard and collapse fragility (i.e. no knowledge of either distribution is needed). However, the second moment approach, as the name implies, provides only the first two moments of the distribution of the collapse probability, and therefore the shape of the distribution must be assumed.

10.7 Semi-parametric Monte Carlo approach

Based on the parametric forms of the seismic hazard and collapse fragility it is possible to use Monte Carlo simulation to generate a non-parametric distribution of the collapse probability for a given time interval. The parametric form of the seismic hazard requires the definition of the mean, epistemic covariance, and epistemic distribution as a function of IM, while the parametric distribution of the collapse fragility curve is likely to be (but not restricted to) the lognormal distribution for both aleatory and epistemic uncertainty [31]. Thus, for simulation *i*, $P_{C,i}$ is obtained by generating realizations of the seismic hazard curve, $P(IM)_i$, and collapse fragility curve, $P(C|IM)_i$, and solving Equation (10-3). By repeating this process *N* times, a total of *N* $P_{C,i}$ values are obtained from which an empirical distribution of P_C can be constructed [28]. This approach has the advantage that full details on the distribution of the ground motion hazard curve may not be available for sites in the San Francisco bay area (as publications generally provide only the mean hazard curve), so one can use the mean hazard curve as given, and the covariance and distribution as investigated in this study to estimate the collapse probability distribution. Also, unlike the second moment method, this approach results in the full collapse probability distribution (i.e. the distribution shape does not need to be assumed).

10.8 Non-parametric logic tree approach

This approach follows directly from the logic-tree approach used to consider epistemic uncertainties in the seismic hazard. The consideration of the epistemic uncertainty in the parameters of the collapse fragility simply represents additional branches on the end of the seismic hazard analysis computation (with say *m* possible options). If the seismic hazard logic tree has *n* end-nodes, then there will be a total of *n*x*m* different values for the collapse probability which can be used to obtain an empirical distribution (as for the semi-parametric Monte Carlo approach above) for the collapse probability. As a result of this continuity, no information is lost by separating the two tasks (seismic hazard and seismic response estimation). This continuity however comes with the likely requirement that the seismic hazard and collapse estimation would have to be conducted for the same site-specific study; as such logic tree details for general sites are not likely to be publicly available.

It should also be noted that epistemic uncertainties in many other variables in the performance-based problem (i.e. structural response, damage and loss) may be represented with continuous distributions rather than the discrete-nature of logic trees. This approach can easily be handled by using Monte Carlo simulation on the end nodes of the logic tree branches and the other continuous random variables, which is in fact desirable even for seismic hazard studies when there are extensive epistemic uncertainties [7].

10.9 Comparison of propagation methods

In order to compare the three different propagation methods described above consider the 30-year collapse probability of a structure located in San Francisco. The (hypothetical) structure has a fundamental period of T_1 = 1.0s and based on seismic response analyses it is determined that the collapse capacity has a lognormal distribution with mean and dispersion of $S_a(T=1.0s) = 1.9g$, and $\sigma_{\ln R,C} = 0.4$, respectively. Due to (epistemic) modelling uncertainties the mean collapse capacity is also uncertain with mean of 1.9g and dispersion $\sigma_{\ln U,C} = 0.4$ (see Zareian and Krawinkler [31] and Haselton [32] for methodological details). No uncertainty is considered in the standard deviation of the collapse capacity (although such higher moment uncertainties can easily be handled in the semi-parametric and non-parametric methods). Unless otherwise noted, Equation (10-2) is used for the correlation structure for the second moment and Monte Carlo approaches, as well as the lognormal assumption for the

epistemic uncertainty in the seismic hazard.

Figure 10-10 illustrates the collapse capacity obtained when considering only epistemic uncertainty in the seismic hazard (i.e. $\sigma_{\ln U,C}$ is zero) using the second-moment method with various assumptions on the correlation structure of epistemic uncertainties in the seismic hazard, and assuming that the collapse probability has a lognormal distribution. Via Equation (10-5) it can be seen that the correlation does not affect the expected value of the 30 year collapse probability, however the covariance terms for the seismic hazard σ_{IM_i,IM_j} , in Equation (10-6) depend on the correlation coefficient and hence give the differing results shown in Figure 10-10. It can be seen that the effect of the correlation is significant with dispersion values of 0.41, 0.77 and 0.84 for none, empirical, and perfect correlation assumptions, respectively.

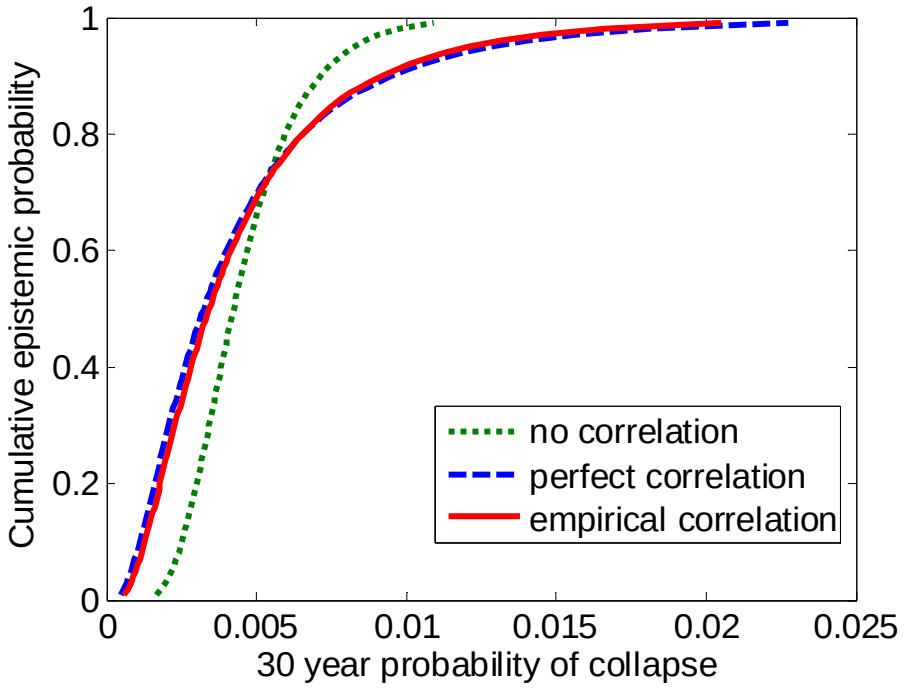


Figure 10-10: Effect of correlation assumption on the distribution of the 30 year probability of collapse.

Figure 10-11a illustrates the distribution of the collapse probability (uncertainty in the seismic hazard only) using the three different propagation methods. As the non-parametric logic tree approach makes no assumptions about the nature of the epistemic uncertainty in the seismic hazard then it can be considered the 'exact' approach. It can be seen that the three methods provide good agreement in the central portion of the distribution with some differences near the tails. This agreement between the cumulative collapse probability distribution is however for a case in which the seismic hazard epistemic uncertainty was well

represented by the lognormal distribution, in situations where this is not the case it is unlikely that the parametric and semi-parametric approaches will produce as similar a result as the non-parametric Logic Tree approach.

Figure 10-11b illustrates the distribution of collapse probability when considering epistemic uncertainty in both seismic hazard and collapse capacity. Again the different methods provide similar results (the second moment and Monte Carlo methods are similar, and thus only one is shown), but more importantly it can be seen that the uncertainty in the collapse probability has been significantly increased when collapse fragility uncertainty is considered. This is in agreement with analytical solutions, which show that the uncertainty in the collapse capacity is $\sigma_{\ln P_C} = \sqrt{k^2 \sigma_{\ln U,C}^2 + \sigma_{\ln U,H}^2}$, where $\sigma_{\ln U,C}^2$ and $\sigma_{\ln U,H}^2$ are the epistemic uncertainty in the collapse capacity and hazard respectively, and k is the log-log slope of the seismic hazard curve (which increases with reducing exceedance probability) [30]. Thus, while it is well acknowledged that there exist large epistemic uncertainties in seismic hazard curves, the current lack of knowledge (i.e. epistemic uncertainty) in collapse prediction appears to be more significant when considering the 30-year collapse probability for the site and structure considered.

Table 2 summarises the pros and cons of the three methods of epistemic uncertainty propagation discussed in this section. While the different methods provide different accuracy, it is most likely to be the input requirements which determine which method is employed (e.g. logic tree details of the seismic hazard must be available to use the non-parametric approach).

Table 10-2: summary of uncertainty propagation methods

Method	Pros	Cons
Second moment (Parametric)	No distribution shape needed for seismic hazard; Computationally efficient.	Requires assumed distribution of performance measure being calculated based on first two moments. Difficulties in handling epistemic uncertainties in higher moments
Monte carlo (Semi-Parametric)	Can consider distribution shape in epistemic uncertainty. Can be used when logic tree details not available.	Computationally more expensive than Second moment method.
Logic tree (Non-Parametric)	Allows direct consideration of the non-parametric form of epistemic uncertainty in seismic hazard.	Requires seismic hazard results from each branch of the logic tree which are not always available.

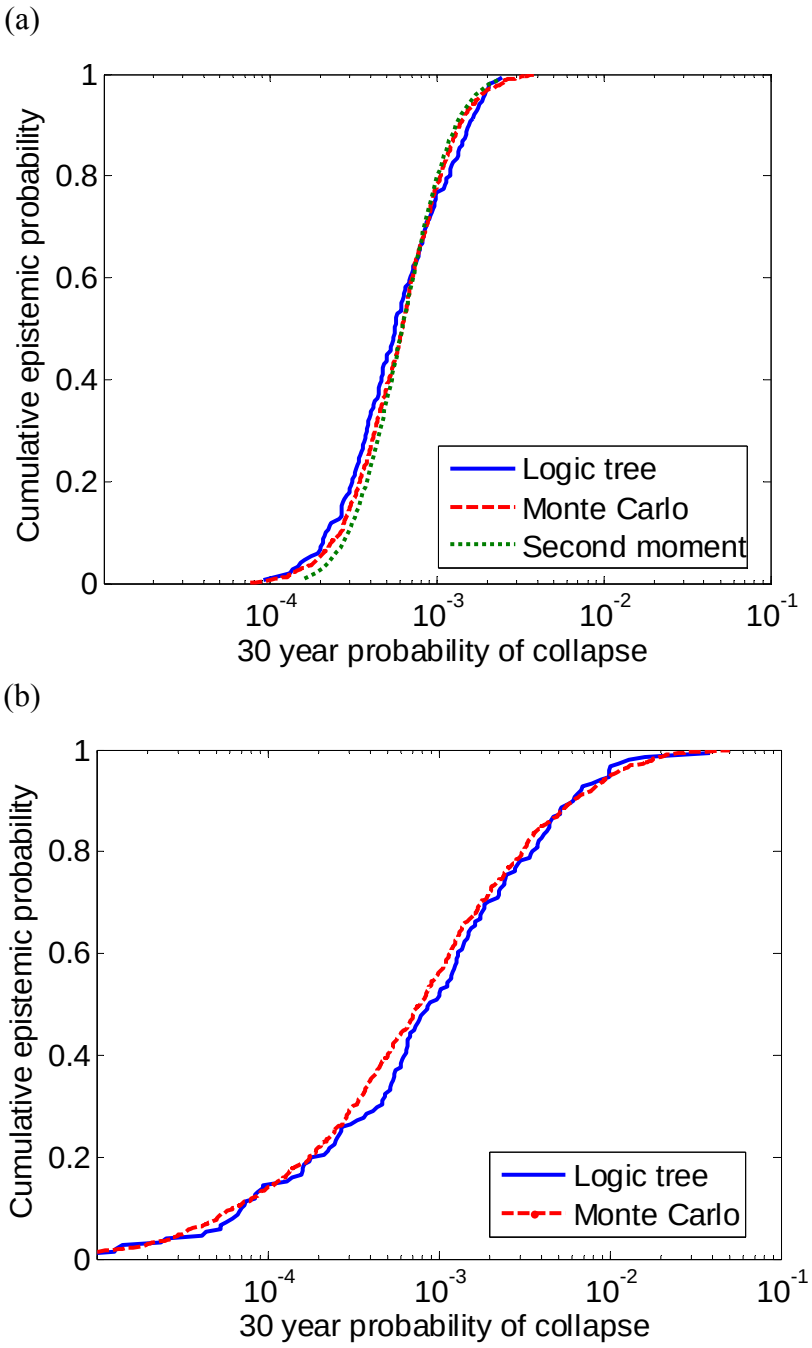


Figure 10-11: Distribution of collapse probability obtained using different uncertainty propagation methods: (a) only seismic hazard epistemic uncertainty; and (b) epistemic uncertainty in both seismic hazard and collapse capacity.

10.10 Conclusions

This chapter has investigated the character of epistemic uncertainty in the results of seismic hazard analyses for various intensity measures and sites in the San Francisco bay area, and their propagation in the seismic performance assessment of structures and facilities. The 2002 Working Group on California Earthquake Probabilities earthquake rupture forecast as

well as two sets of ground motion prediction equations were used to rigorously capture the epistemic uncertainty in the seismic hazard for the different intensity measures and sites investigated. The magnitude of the epistemic uncertainty was observed to be significant and increased with reducing probability of exceedance. With appropriate normalisation it was illustrated that the magnitude of the epistemic uncertainty can be approximately estimated as a function of the mean probability of exceedance. For the specific sites and intensity measures considered the similarity in the ground motion prediction equations was observed to have a strong influence on the magnitude of the epistemic uncertainty.

In the performance-based assessment of structures and facilities it maybe necessary to have knowledge of the distribution and correlation structure of the epistemic uncertainty. In situations in which the epistemic uncertainty in ground motion prediction was the same order as the epistemic uncertainty in the earthquake rupture forecast, the total seismic hazard epistemic uncertainty was well approximated by the lognormal distribution. When the epistemic uncertainty in ground motion prediction was notably larger than the epistemic uncertainty in the earthquake rupture forecast, the epistemic uncertainty in the seismic hazard was generally not lognormally distributed. The correlation between the epistemic uncertainty in the probability of exceedance at two different intensities was observed to be a function of the ratio of the two intensity levels and a simple equation was proposed to predict this correlation.

Propagation of seismic hazard epistemic uncertainty to estimate the epistemic uncertainty in seismic performance assessment can be addressed in several ways of varying complexity and input requirements. Comparisons between three methods for the 30-year collapse probability of a hypothetical structure illustrated that the methods yielded similar results. The epistemic uncertainty in the collapse probability of the hypothetical structure was observed the dominant contributor to the uncertainty in 30-year collapse probability compared to epistemic uncertainty in seismic hazard curves.

The observations made regarding the magnitude, distribution and correlation of the epistemic uncertainties in seismic hazard analyses are specific to the sites and intensity measures considered and may not be appropriate for use in other tectonic regions or for other intensity measures.

10.11 Acknowledgements

Financial support from the New Zealand Tertiary Education Commission is appreciated.

Dr Ned Field, Mr Kevin Milner, and Dr Christine Goulet are greatly thanked for their roles in developing the OpenSHA applications used in this study. Comments from two anonymous reviewers are also gratefully acknowledged.

10.12 References

[1] Der Kiureghian A and Ditlevsen O. Aleatory or epistemic? Does it matter? Structural Safety 2008; 31(2): 105-112, DOI: 10.1016/j.strusafe.2008.06.020.

[2] Der Kiureghian A. Non-erogodicity and PEER's framework formula. Earthquake Engineering and Structural Dynamics 2005; 34(13): 1643-1652, DOI: 10.1002/eqe.504.

[3] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. PEER Center News 2000; 3(2).

[4] Baker JW and Cornell CA. Uncertainty propagation in probabilistic seismic loss estimation. Structural Safety 2008; 30(3): 236-252.

[5] Kramer SL. Geotechnical Earthquake Engineering. Prentice-Hall: Upper Saddle River, NJ., 1996; 653.

[6] McGuire RK. Seismic Hazard and Risk Analysis. Earthquake Engineering Research Institute, 2004; 221.

[7] WGCEP. Earthquake Probabilities in the San Francisco Bay Region: 2002–2031. USGS Open-File Report 03-214, 2003.

[8] Kulkarni RB, Youngs RR, and Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. 8th World Conference on Earthquake Engineering,, San Francisco, CA, 1984.

[9] Field EH, Gupta N, Gupta V, Blanpied M, Maechling P, and Jordan TH. Hazard calculations for the WGCEP-2002 forecast using OpenSHA and distributed object technologies. Seismological Research Letters 2005; 76: 161-167.

[10] Field EH, Jordan TH, and Cornell CA. OpenSHA: A Developing Community-Modelling Environment for Seismic Hazard Analysis. Seismological Research Letters 2003; 74: 406-419.

[11] Campbell KW and Bozorgnia Y. Updated near-source ground motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra. Bulletin of the Seismological Society of America 2003; 93(1): 314-331.

[12] Frankel AD, Petersen MD, Mueller CS, Haller KM, Wheeler RL, Leyendecker EV, Wesson RL, Harmsen SC, Cramer CH, Perkins DM, and Rukstales KS. Documentation for the 2002 Update of the National Seismic Hazard Maps. United States Geological Survey 2002.

[13] Boore DM, Joyner WB, and Fumal TE. Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work. Seismological Research Letters 1997; 68(1): 128-153.

[14] Abrahamson NA and Silva WJ. Empirical response spectral attenuation relations for shallow crustal earthquakes. Seismological Research Letters 1997; 68(1): 94–126.

[15] Sadigh K, Chang CY, Egan JA, Makdisi F, and Youngs RR. Attenuation relationships for shallow crustal earthquakes based on California strong motion data. Seismological Research Letters 1997; 68(1): 180-189.

[16] Campbell KW and Bozorgnia Y. Campbell-Bozorgnia NGA horizontal ground motion model for PGA, PGV, PGD and 5% damped linear elastic response spectra. Earthquake Spectra 2008; 24(1): 139-171.

[17] Boore DM and Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s. Earthquake Spectra 2008; 24(1): 99-138.

[18] Abrahamson NA and Silva WJ. Summary of the Abrahamson & Silva NGA ground motion relations. Earthquake Spectra 2008; 24(1): 67-97.

[19] Chiou BSJ and Youngs RR. An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. Earthquake Spectra 2008; 24(1): 173-215.

[20] Idriss IM. An NGA Empirical Model for Estimating the Horizontal Spectral Values Generated By Shallow Crustal Earthquakes. Earthquake Spectra 2008; 24(1): 217-242.

[21] Abrahamson NA, Atkinson GM, Boore DM, Bozorgnia Y, Campbell KW, Chiou B, Idriss IM, Silva WJ, and Youngs RR. Comparisons of the NGA Ground-Motion Relations. Earthquake Spectra 2008; 24(1): 45-66.

[22] Campbell KW and Bozorgnia Y. Campbell-Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters. University of California, Berkeley, 2007. 238.

[23] Abrahamson NA. Seismic hazard assessment: problems with current practice and future developments. 1st European Conference on Earthquake Engineering and Seismology, Geneva, Switzerland 2006.

[24] Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, and Rukstales KS. Documentation for the 2008 Update of the United States National Seismic Hazard Maps. United States Geological Survey (USGS), 2008. 127.

[25] Abrahamson NA and Bommer JJ. Probability and uncertainty in seismic hazard analysis. Earthquake Spectra 2005; 21(2): 603-607.

[26] Musson RMW. Against Fractiles. Earthquake Spectra 2005; 21(3): 887-891.

[27] McGuire RK, Cornell CA, and Toro GR. The Case for Mean Seismic Hazard. Earthquake Spectra 2005; 21(3): 879-886.

[28] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[29] Scherbaum F, Bommer JJ, Bungum H, Cotton F, and Abrahamson NA. Composite Ground-Motion Models and Logic Trees: Methodology, Sensitivities, and Uncertainties. Bulletin of the Seismological Society of America 2005; 95(5): 1575-1593, DOI: 10.1785/0120040229.

[30] Bradley BA and Dhakal RP. Error estimation of closed-form solution for annual rate of structural collapse. Earthquake Engineering and Structural Dynamics 2008; 37(15): 1721-1737.

[31] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1901-1914.

[32] Haselton CB. Assessing Collapse Safety of Modern Reinforced Concrete Moment Frame Buidlings. Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, 2007, 312.

11.Epistemic Uncertainties in Component Fragility Functions

Bradley BA. Epistemic uncertainties in component fragility functions. *Earthquake Spectra* 2009. (in press)

11.1 Abstract

This chapter is concerned with the inclusion of epistemic uncertainties in component fragility functions used in performance-based earthquake engineering. Conventionally fragility functions, defining the probability of incurring at least a specified level of damage for a given level of seismic demand, are defined by a mean and standard deviation and assumed to have a lognormal distribution. However, there exist many uncertainties in the development of such fragility functions. The sources of epistemic uncertainty in fragility functions, their consideration, combination, and propagation are presented and discussed. Two empirical fragility functions presented in literature are used to illustrate the epistemic uncertainty in the fragility function parameters due to the finite size of the datasets. These examples and the associated discussions illustrate that the magnitude of epistemic uncertainties are significant and there are clear benefits of the consideration of epistemic uncertainties pertaining to the documentation, quality assurance, implementation, and updating of fragility functions. Epistemic uncertainties should therefore always be addressed in future fragility functions developed for use in seismic performance assessment.

11.2 Introduction

Emerging trends in earthquake engineering involve the design and assessment of seismic performance based on structural response at multiple levels of ground motion intensity, and even the use of loss estimation methods for explicit seismic performance quantification of entire facilities. Central to these trends is an understanding of how earthquake-induced dynamic response is related to seismic performance in terms of the consequences of the associated damage. The overall performance of a facility will be a function of the seismic performance of all of its components which will include structural and non-structural components as well as facility contents. Hence, relationships are needed in order to make the link between seismic demand and damage to the components of a facility, which are commonly referred to as seismic fragility functions. As this chapter is concerned solely with seismic excitations, for brevity the term 'seismic' is omitted herein when referring to seismic fragility functions.

Because of the discrete nature of post-earthquake repair of damaged components and the many uncertainties associated with the prediction of seismic performance, fragility functions provide the probability of incurring, or exceeding, a specified discrete damage state (DS) as a function of some engineering demand parameter (EDP). Fragility functions can be used to represent a structure as a whole [1], or to represent individual structural [2, 3], non-structural [4, 5], and contents [6, 7] components within a facility. This chapter is focused solely on component fragility functions, but many of the discussions also apply to the former.

Such component fragility functions are typically determined using one or a combination of experimental, numerical, and observational data, as well as expert opinion. Recognising the importance of fragility functions in the bigger picture of performance-based earthquake engineering Porter *et al*. [8], as part of the ATC-58 project [9], provide standardisation for the development and documentation of fragility functions. FEMA 461 [10] also provides guidelines on the development of fragility functions based on data from laboratory experiments. As fragility functions are typically assumed to have lognormal distributions, their conventional development involves estimation of the mean and standard deviation that uniquely define the lognormal distribution.

Uncertainties in earthquake engineering are typically classified as being either aleatory or epistemic. Apparently random observations due to: unknown factors; known factors which are not modelled; and pragmatic simplifications of reality, are generally considered to be a

result of aleatory uncertainties [11]. Epistemic uncertainties, on the other hand, are due to a lack-of-knowledge and can, in theory, be reduced with improved understanding of the problem. While the distinction of aleatory and epistemic uncertainties is somewhat problem-dependent, it is prudent to classify such uncertainties as either aleatory or epistemic based on those which can and cannot be reduced by the analyst [12]. In the consideration of time-dependent problems it is also important to understand that epistemic uncertainties are typically non-ergodic, while aleatory uncertainties are ergodic by definition [12, 13].

As conventional fragility functions are typically based on estimates of the distribution parameters from data alone, then the standard deviation of this data represents only aleatory uncertainty. It should seem obvious however that there can be significant epistemic uncertainty in the estimated parameters due to the imperfect data and/or knowledge used in their estimation. Such epistemic uncertainty should also be considered, whether it is combined with the aleatory uncertainty or is treated separately. It is both strictly correct and also beneficial to consider such epistemic uncertainties in the development of fragility functions irrespective of how they are propagated in their application.

Owing to a lack of empirical data at the time, Kennedy *et al*. [14] developed estimates of aleatory and epistemic uncertainties in fragility functions for the Oyster Creek nuclear power plant from "analysis combined heavily with engineering judgment" and discussed how such uncertainties can be either treated separately or combined as a composite uncertainty. The purpose of this chapter is to address epistemic uncertainties which arise in the acquisition and manipulation of data used to develop component fragility functions. Following an overview of fragility functions and how they can be obtained, the influence of fragility function uncertainty in seismic performance is addressed. Various epistemic uncertainties in fragility functions are then discussed and methods to address and rigorously combine them are outlined. Two examples of fragility functions published in literature are then used to illustrate the significance and treatment of these epistemic uncertainties.

11.3 Overview of fragility functions

Seismic fragility functions are typically used to describe the uncertainty in the capacity of a component which is subjected to seismic demand. More specifically, fragility functions give the probability of being in or exceeding a specific damage state (DS) as a function of the component engineering demand parameter (EDP), i.e. $F(DS \geq ds|EDP = edp)$. Because of its simplicity and apparent applicability, the lognormal distribution is typically used to define the

fragility function [8]:

$$F(ds_i \mid edp) = \Phi\left(\frac{\ln edp - \mu_{\ln EDP|DS_i}}{\sigma_{\ln EDP|DS_i}}\right) \tag{11-1}$$

where $\Phi(\ )$ is the cumulative normal distribution variate; and $\mu_{\ln EDP|DS_i}$ and $\sigma_{\ln EDP|DS_i}$ are the mean and standard deviation of $lnEDP|DS_i$, i.e. the capacity. Thus, from the above formulation it can be seen that only the mean and standard deviation (or the variance) of the capacity are required to define the fragility function.

Components may have multiple damage states, characterised by the discrete levels of repair required to restore the component to its undamaged state. In such cases each damage state is defined using a fragility function. As component fragility functions are generally *sequential* (i.e. the occurrence of $DS = ds_i$ implies the occurrence of $DS = ds_{i-1}$) and as the fragility function gives the probability of being equal to or larger than a specified level of damage, then the probability of being *in* a specific damage state is given by:

$$P(ds_i \mid edp) = \begin{cases} F(ds_i \mid edp) - F(ds_{i+1} \mid edp) & 1 \le i < N_{DS} \\ F(ds_i \mid edp) & i = N_{DS} \end{cases} \tag{11-2}$$

where N_{DS} is the number of damage states. Figure 11-1 illustrates a typical set of fragility functions for an reinforced concrete (RC) beam [15]. It is noted that for a given EDP, the damage state probabilities given by Equation (11-2) represent the vertical difference between the fragility functions.

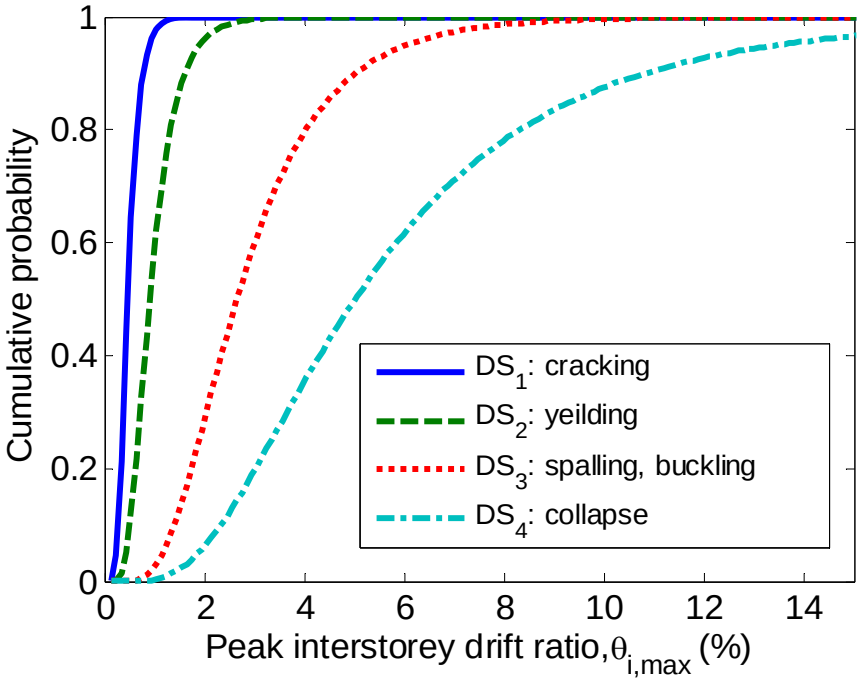


Figure 11-1: Fragility functions for the four damage states in a RC column [15].

11.4 Methods of developing fragility functions

Before discussing the causes of uncertainties in the fragility functions, and methods to compute such uncertainties, it is necessary to first briefly review various methods which can be used to develop fragility functions. Porter *et al*. [8] discusses various methods to compute fragility functions, and a brief summary is given here for completeness. The five different methods for computing fragility functions given in Porter *et al*. [8] are: (A) Actual EDP; (B) Bounding EDP; (C) Capable EDP; (D) Derived i.e. analytical; (E) Expert opinion; and (U) updating.

Method A: Actual EDP provides the EDP value at which failure of the specimen occurred. This level of information is typically only possible in controlled laboratory experiments [2, 3]. The parameters of the lognormal distribution can be determined either from the method of moments, maximum likelihood estimation, or by using generalised linear regression with a *probit* link and the empirical distribution function [16]. The validity of the lognormal distribution can be assessed using the Lilliefors test statistic [17].

Method B: Bounding EDP provides the peak EDP value the component was subjected to and whether or not failure (i.e. damage state exceedance) was observed. Thus, as opposed to method A which provides continuous EDP data, method B provides binary EDP data (i.e. failure did or did not occur). Such data is most likely obtained from post-earthquake

reconnaissance surveys [5, 18]). Porter *et al*. [8] discuss one method based on binning of data that can be used, however it is the authors opinion that the distribution parameters can be easily and more rigorously obtained by performing logistic regression [16]. The use of logistic regression, as opposed to conventional least-squares regression, is necessary to account for the non-constant (heteroskedastic) variance in the regression, i.e. that an error of 0.1 between the cumulative distribution function (CDF) and the data of is more significant when the CDF value is 0.88 and the data 0.98, compared to a CDF value of 0.45 and data 0.55.

Methods C, D, and E provide various means to use analysis and/or apply subjective judgement in the absence of sufficient empirical data on component failure. Method U provides a means to update a fragility function by using additional data which was not available or not used at the time the fragility function was first created.

The remainder of this chapter is primarily concerned with the epistemic uncertainties when using methods A and B as outlined above. That said, many of the comments made are equally applicable to the other methods.

11.5 Uncertainty in damage state probability

As discussed in this chapter, there exist epistemic uncertainties in the determination of the distribution parameters of fragility functions. As a result, there will be uncertainty in damage states probabilities given in Equation (11-2). The following section discusses how, given uncertainties in the distribution parameters, such damage state probability uncertainties can be determined.

11.5.1 Uncertainty in damage state probability for a given intensity

In an assessment of seismic performance, fragility functions can be convolved with the seismic response to determine the probability of being *in* a specified damage state for a given level of a ground motion intensity measure (IM):

$$\begin{aligned} P\left(ds_i|im\right) &= \int_{EDP} P\left(ds_i|edp\right) f\left(edp|im\right) dEDP \\ &\cong \sum_j w_j P\left(ds_i|edp_j\right) f\left(edp_j|im\right) \end{aligned} \tag{11-3}$$

where *IM* is the intensity measure that is used to quantify the intensity of the ground motion; $P\left(ds_i|im\right)$ is the probability of incurring $DS = ds_i$ given $IM = im$; $P\left(ds_i|edp\right)$ is the

probability of being *in DS* = ds_i given EDP = *edp* (i.e. Equation (11-2)); $f(edp|im)$ is the probability density function for the seismic response, *EDP* = *edp*, conditioned on *IM* = *im*; and w_j is the numerical integration weight [19].

When epistemic uncertainties are present, both $P(ds_i|edp)$ and $f(edp|im)$ are uncertain variables. There are various methods by which uncertainties in the arguments of Equation (11-3) can be propagated to determine the uncertainty in $P(ds_i|im)$ (e.g. simulation or analytical methods). Here statistical moments (i.e. mean and variance) are used to illustrate the features of $P(ds_i|edp)$ and $f(edp|im)$ which need to be considered, and it is left to the individual to select the uncertainty propagation method for their particular application. The expectation and variance of the discrete form of Equation (11-3) can be shown to be (i.e. Appendix C):

$$E\left[P(ds_i|im)\right] \cong \sum_j w_j \mu_{X_j} \mu_{Y_j} \tag{11-4}$$

$$Var\left[P(ds_i|im)\right] = \sum_j \sum_k \left\{ w_j w_k \left(\sigma_{X_j,X_k} \sigma_{Y_j,Y_k} + \mu_{X_j} \mu_{X_k} \sigma_{Y_j,Y_k} + \mu_{Y_j} \mu_{Y_k} \sigma_{X_j,X_k} \right) \right\} \tag{11-5}$$

where $X_j = P(ds_i|edp_j)$; $Y_j = f(edp_j|im)$; $\mu_{X_j} = E[X_j]$ is the expectation of X_j; and $\sigma_{X_j,X_k} = Cov[X_j, X_k]$ is the covariance of X_j and X_k. From Equations (11-4) and (11-5) it can be seen that in order to determine the distribution of $P(ds_i|im)$, the mean and covariance of the uncertainty in the damage states for a given component is required. The focus in this chapter is the determination of the epistemic uncertainties in $X_j = P(ds_i|edp_j)$, while various studies have investigated the epistemic uncertainty in the seismic response distribution, $f(edp|im)$ [e.g. 20].

11.5.2 Uncertainty in damage state probability given *EDP*

Because of the various epistemic uncertainties in fragility functions, for a given level of demand, there exists an uncertain value of the probability of being *in* a given DS due to the imposed demand, EDP. As the probability of being *in* a given damage state is given by:

$$P(ds_i \mid edp) = F(ds_i \mid edp) - F(ds_{i+1} \mid edp) \tag{11-6}$$

where the case of $i = N_{DS}$ in Equation (11-2) can be handled by defining $F\left(ds_{N_{DS}+1} \mid edp\right) = 0$, then the expected value and variance in the probability of being in $DS = ds_i$ given $EDP = edp$ is given by:

$$E\left[P\left(ds_i \mid edp\right)\right] = E\left[F\left(ds_i \mid edp\right)\right] - E\left[F\left(ds_{i+1} \mid edp\right)\right] \tag{11-7}$$

$$\begin{aligned} Var\left[P\left(ds_i \mid edp\right)\right] &= E\left[\left\{P\left(ds_i \mid edp\right)\right\}^2\right] - E\left[\left\{P\left(ds_i \mid edp\right)\right\}\right]^2 \\ &= Var\left[F_i\right] + Var\left[F_{i+1}\right] - 2Cov\left[F_i, F_{i+1}\right] \end{aligned} \tag{11-8}$$

where $F_i = F\left(ds_i \mid edp\right)$; $Var\left[F_i\right]$ is the variance in F_i; and $Cov\left[F_i, F_{i+1}\right]$ is the covariance between F_i and F_{i+1}. Similarly it can be shown that the covariance in the damage state probabilities is given by:

$$\begin{aligned} Cov\left[P\left(ds_i \middle| edp\right), P\left(ds_j \middle| edp\right)\right] &= Cov\left[F_i, F_j\right] - Cov\left[F_i, F_{j+1}\right] \\ &\quad - Cov\left[F_{i+1}, F_j\right] + Cov\left[F_{i+1}, F_{j+1}\right] \end{aligned} \tag{11-9}$$

Thus, determination of the epistemic uncertainty in damage states requires both knowledge of the mean and variance in the damage state fragility curves (i.e. uncertainty in the F_i's, due to epistemic uncertainty in the distribution parameters, $\mu_{\ln EDP|DS_i}$ and $\sigma_{\ln EDP|DS_i}$), as well as the covariance (or correlation) between the different damage state fragility curves.

11.5.3 Combining aleatory and epistemic uncertainties

Here it is timely to introduce the simplest method for the incorporation of epistemic uncertainties into a fragility function. While the majority of this chapter is dedicated to the separate consideration and propagation of aleatory and epistemic uncertainties, it is also possible to combine both uncertainties and treat them as a composite uncertainty. If it is assumed that the uncertainty in the (logarithmic) mean of the component fragility function is normally distributed and there is no uncertainty in the variance of the fragility function, then the mean value (with respect to epistemic uncertainties) of the fragility function is also given by a lognormal distribution with the same mean, $\mu_{\ln EDP|DS_i}$, and standard deviation given by [14]:

$$\sigma_{\ln EDP|DS_i} = \sqrt{\sigma^2_{\ln EDP|DS_i(a)} + \sigma^2_{\ln EDP|DS_i(e)}} \tag{11-10}$$

where $\sigma_{\ln EDP|DS_i(a)}$ and $\sigma_{\ln EDP|DS_i(e)}$ are the lognormal standard deviations due to aleatory and epistemic uncertainty, respectively, and are assumed to be independent. Because of its simplicity this 'square-root-sum-of-the-squares' (SRSS) method has been used widely in

earthquake engineering, for example, Cornell *et al*. [21].

11.6 Epistemic uncertainty in fragility functions

In this section various sources of epistemic uncertainties are introduced as well as methods to determine the magnitude of several of these uncertainties. Examples are provided in subsequent sections.

The primary cause of epistemic uncertainties in fragility functions determined using methods A and B is that of limited high quality data. Firstly, this means that the parameters defining the fragility function (i.e. the median and dispersion) are determined based on a finite sample of data that, although representative, is a random realisation of the true population. This uncertainty is referred to as *finite sample uncertainty* and, as will be shown, is related to the number of independent data points used to determine the fragility functions. Secondly, one aspect of data quality refers to how well the conditions under which the data were obtained simulate the conditions in which the component exists in reality. This uncertainty is referred to as *in-situ uncertainty*. Another aspect of data quality is the method in which the seismic demand is imposed on the specimen, and is referred to as *loading protocol uncertainty*. With particular reference to methods C, D, and E there will also be further uncertainties due to lack of knowledge.

Table 11-1 provides seven sources of uncertainty present in the development of fragility functions from experimental data as defined in FEMA 461 [10]. Each of the sources has been classified as one of the aforementioned epistemic uncertainties, while the last two uncertainties in Table 11-1 are simply denoted as *randomness* (e.g. aleatory) uncertainty which can be directly considered using the experimental data. As noted in FEMA 461 [10], the first four, *in-situ*, uncertainties can be reduced by careful planning of the experimental program. While the uncertainties listed in Table 11-1 were specifically for experimental data (i.e. typically method A), they also apply to those based on empirical data from post-earthquake reconnaissance (i.e. typically method B).

Table 11-1: Sources of uncertainty in fragility functions as identified in FEMA461 [10].

N	Description	Classification
1	Testing a component isolated from its in-situ conditions such as electrical conduits, piping, or supported floor slabs	*in-situ*
2	Imperfect simulation of boundary conditions	*in-situ*
3	Extrapolation to in-situ conditions not fully simulated in the test	*in-situ*
4	Variability in configuration	*in-situ*
5	Employment of a loading history that cannot precisely replicate the loading experienced by components in a real building responding to earthquake shaking	*loading protocol*
6	Uncertainty in the definition of the several damage states, and the input loading at which they initiate	*randomness*
7	Variability in material properties and fabrication/construction methods and details	*randomness*

11.6.1 Finite sample uncertainty

Finite sample uncertainty occurs because the distribution (defined in this case by the mean and standard deviation) of the population is estimated based on statistics of a sample. It should be noted that this source of uncertainty is not listed in Table 11-1. Two methods are discussed here to characterise the uncertainty in the distribution parameters based on *finite sample uncertainty*.

In the case of fragility function development using method A, the *N* independent and identically distributed data can be used to determine the sample mean and sample variance of the logarithm of the data; i.e. $\overline{X}_{\ln EDP|DS_i}$ and $S^2_{\ln EDP|DS_i}$, respectively. Assuming that the (logarithm of the) data are obtained from a normal distribution it can be shown that the mean and variance in the sample mean and sample variance are given by Equations (11-11) and (11-12), respectively [22, pg. 255-257]:

$$E\left[\overline{X}_{\ln EDP|DS_i}\right] = \mu \quad ; \qquad Var\left[\overline{X}_{\ln EDP|DS_i}\right] = \frac{\sigma^2}{N} \tag{11-11}$$

$$E\left[S^2_{\ln EDP|DS_i}\right] = \sigma^2 \quad ; \qquad Var\left[S^2_{\ln EDP|DS_i}\right] = \frac{2\left(\sigma^2\right)^2}{N-1} \tag{11-12}$$

From Equations (11-11) and (11-12) it can be seen that the variance in both the sample mean and sample variance are proportional to *1/N*, giving the intuitive result that the estimate of these distribution parameters becomes less uncertain as the sample size increases. It can also be shown that the sample mean and sample variance (once appropriately normalised) have Student's-*t* and Chi-Square distributions, both of which as the sample size increases

approach the normal distribution [22, pg. 255-257].

As opposed to the use of analytical methods to determine the uncertainty in distribution parameters, sampling methods are also possible with the most versatile being bootstrap sampling [23]. Bootstrap sampling involves sampling with replacement from the finite dataset. For each bootstrap sample the desired statistics (in this case the mean and variance) are computed. This process is then repeated *N* times, giving *N* different values of the sample mean and sample variance from which their statistics can be obtained. In the limit as the number of bootstrap samples tends to infinity (although approximately 500 provides more than ample accuracy and is used herein) the results tend toward those obtained analytically (e.g. Equations (11-11) and (11-12)). In addition, bootstrap sampling can: (i) also be used to handle data of the 'method B' type, for which logistic regression is required to determine the distribution parameters; and (ii) it can be used to combine other epistemic uncertainties as discussed in the following sections.

It is important to note that if an *unbiased* estimate of the sample variance [22, pg. 257] is computed using either the analytical approach (e.g. Equations (11-11) and (11-12)) or bootstrap simulation, that this unbiased estimate already accounts for the uncertainty in the estimate of the sample mean, and thus should not be 'double-counted' if it is decided to combine aleatory and epistemic uncertainties using Equation (11-10).

Equation (11-11) can be used as a guide to the number of replicates (i.e. data) that is required in order to achieve a specified level of uncertainty in the mean value of the damage state capacity. For example, FEMA 461 [10] notes that "a minimum of 3 specimens should be used to determine the fragility function of the component". If, for the sake of argument, the sample logarithmic mean is $E\left[\overline{X}_{\ln EDP|DS_i}\right]=0.0$ (equivalent to a sample median of $\ln(0.0)=1.0$) and sample standard deviation is $E\left[S^2_{\ln EDP|DS_i}\right]=0.4$ (the reference value suggested in Porter *et al*. [8]), then using Equations (11-11) and (11-12), the 90% confidence interval for the sample logarithmic mean and variance are [22, pg. 263, 269]:

$$\left\langle \overline{X}_{\ln EDP|DS_i} \right\rangle_{0.90} = \left(\ln(0.51), \ln(1.96)\right) \tag{11-13}$$

$$\left\langle S^2_{\ln EDP|DS_i} \right\rangle_{0.90} = \left(0.23^2, 1.77^2\right) \tag{11-14}$$

where $\langle\ \rangle_{0.90}$ represents a 90% confidence interval of the argument. Clearly, the ranges of values in the above confidence intervals are significant and such uncertainties should be understood by developers and user's of fragility functions.

11.6.2 Loading protocol: Laboratory-based data

In assessing *loading protocol uncertainty* it is beneficial to distinguish between the three common forms of seismic testing: dynamic 'shake table'; quasi-static; and pseudo-dynamic. While shake table testing is identified as the most direct method of assessing seismic performance, the capabilities of the shake table apparatus can impose severe limitations, and specimens are commonly tested at reduced size. Such scale effects can be important as many materials and components do not strictly follow similitude rules [24, 25]. At the other end of the spectrum, the nature of quasi-static testing means that larger specimens can be commonly tested. Quasi-static testing, however, does not consider strain rate effects, typically resulting in lower strength and stiffness than would be obtained in dynamic testing [25]. In regard to fragility functions for minor damage states, strain rate effects are also particularly important in the extent of crack propagation, for which quasi-static testing will yield conservative results [25]. Contemporary pseudo-dynamic testing performed on full scale sub-assemblages in real time, and advances in full scale shake-table facilities offer significant reductions in bias and uncertainty in regard to the scale and strain rate effects mentioned above. If sub assemblages, as opposed to entire structures, are tested (typically using either quasi-static or pseudo-dynamic methods) there are also various typical assumptions such as constant axial loads, constant shear-to-moment ratios; and boundary conditions (e.g. fixed inflexion points) [24].

There are various other uncertainties in the acquisition of experimental data in addition to those uncertainties associated with the type of experimental method adopted. Such uncertainties are generally the result of shortfalls in the three key aspects of experimental research: *replication*, *randomisation*, and *blocking* [26]. *Replication* involves repetition of the same experiment arrangement with different specimens, and is necessary to account for experimental errors and random uncertainties associated with the particular design. Thus the uncertainty associated with *replication* is that of *finite sample uncertainty* addressed in the previous section. *Randomisation* involves the random sequencing of tests and their material, geometrical and loading characteristics. Geometrical *randomisation* could include the effects of different boundary conditions, while loading characteristics are particularly important and are discussed in the following paragraph. *Blocking* is a design technique used to improve the precision with which comparisons of several factors are made. This has less relevance in the case of destructive testing and is not discussed further.

The effects of experimental loading history are important as component capacity depends on several aspects of the seismic demand, in particular, the number, magnitude and

sequence of inelastic excursions [27]. Krawinkler [27] discusses the features which influence the character of these aspects of seismic demand, primarily the vibration period of the structure and earthquake magnitude. Because of this significant variability, Krawinkler [27] suggests that for generic components the need exists to use loading protocols which represent demands for short period structures due to large magnitude events, with the understanding that they may be very conservative for long period structures and/or small magnitude events. Some of this uncertainty associated with the ground motion is aleatory uncertainty (i.e. that due to earthquake magnitude), while some is epistemic uncertainty and non-ergodic in time (i.e. that due to the period of vibration of the structure). The current approach mentioned in FEMA 461 [10] is that the component fragility is assumed independent of the structure and ground motion because "such effects are not predictable" (page 5) and that they are necessary "in order to make the problem tractable" (page 40). To some extent (i.e. the aleatory component) this statement is correct, but this does not mean to say that it should not be considered as additional uncertainty in developing fragility functions. Obviously one method to directly consider the effects of loading history is to use different loading histories in the experiments (i.e. different ground motions for shake-table or pseudo-dynamic tests, or different quasi-static loading regimes) for different replicates (i.e. this is *randomisation* of the loading history).

Clearly *randomisation* of loading history will result in larger uncertainty in the component damage state capacity than if the same loading history was used in all replicates. This uncertainty can perhaps be reduced by using an EDP which accounts for both the peak and cumulative nature of the seismic demand, as opposed to the common EDP's of peak displacement (or peak interstorey drift) and peak floor acceleration. Examples of such cumulative EDP's can be found in the state-of-the-art review of damage indices by Williams and Sexsmith [28]. In comparison to the acknowledged effect of loading history on structural performance, its significance for non-structural components is, as far as the author is aware, unknown.

The method by which measurements are made in an experiment can also cause uncertainties. Because of the uncertain definition of damage states, the occurrence of minor damage is typically based on the visual appearance of the component being tested. It is conventional in experimental research to only closely inspect specimens for damage (cracking etc.) at the peak values of loading cycles, therefore there is uncertainty as to the actual EDP value at which the specified damage state was reached [2]. This is referred to as *drift increment uncertainty* herein.

11.6.3 Loading protocol: Post-earthquake reconnaissance data

Data obtained from post-earthquake reconnaissance explicitly considers many of the deficiencies of current laboratory experimental methods identified in the previous section. Again, the key aspects of *replication* and *randomization* are important for this form of data. That is, ideally there should be a large number of data available which are largely independent (i.e. random) of each other. Typical examples of dependence among observations include the similarity in the ground motion experienced by two different components (i.e. they are in the same structure); observations of the same component during multiple earthquakes; and the similarity in the in-situ conditions of two components (i.e. installed by the same contractor in a building). Such dependencies violate the assumption that the individual data are independent and identically distributed, an assumption which is used when determining the parameters of the lognormal distribution defining the fragility function. There are various methods which can be employed to account for such dependencies in observed data when performing regression, and can be found in Straub and Der Kiureghian [29] and references therein.

It was previously mentioned that in the case of data obtained using laboratory experiments, the aleatory uncertainty in the fragility of the component can be reduced if the EDP accounts for both the peak and cumulative demands. However, the lack of data regarding the input seismic demand is likely to mean that, at best, peak ground motion measures are available [5]. An exception to the previous comment would be a structure with a dense deployment of sensors and other seismic instrumentation.

11.6.4 Host-to-target uncertainty

In addition to the epistemic uncertainties in the development of fragility functions there are also inevitably uncertainties in their application. The key uncertainty in this regard is referred to as *host-to-target* uncertainty, and represents the difference in those components whose data provide the basis for the fragility function development (the *host*) with that of the component in the structure of interest for which a fragility function is desired (the *target*). In many cases this uncertainty may be zero, the *target* component is identical (in theory) to the *host* components used to create the fragility function. However, in certain cases fragility functions for a particular *target* component may not be available and it may be infeasible to perform component-specific testing to determine necessary fragility functions. Fragility functions for a similar *host* component may however be available. For example, the *host*

component could be a similar component used in another country that has the same purpose, but composed of slightly different materials and/or having different installation techniques. Such uncertainty due to this *host-to-target* conversion should be considered when applying fragility functions, the magnitude of which will likely be determined by expert judgement on a case-by-case basis.

11.6.5 Needs in uncertainty specification and quantification for laboratory experiments

As previously mentioned there are several uncertainties in laboratory experiments which may be significant, but are not quantifiable from the data obtained for developing the fragility function alone. In such cases, default values are needed based on current knowledge (most likely decided by a panel of experts). Examples of such uncertainties from those discussed previously include: the bias and uncertainty in cracking and other minor damage states due to the neglection of strain rate effects; uncertainties in experimental boundary conditions; and uncertainty due to random loading history as opposed to a single code-based quasi-static protocol. If future experiments are conducted with some eventual intent on developing fragility functions then it is probably not overly difficult to account for some of these uncertainties in the experimental program. The development of fragility functions by compiling experimental data obtained by others is likely to result in the aforementioned uncertainties since such experiments were not conducted with some of the above points in mind.

11.7 Correlations between fragility function parameters

11.7.1 Causes of epistemic correlations

Equation (11-8) indicates that epistemic uncertainty in the probability of being *in* a specific DS given EDP depends not only on the epistemic uncertainty in the parameters of the two fragility functions defining the boundary of that damage state but also on the correlation between the uncertain parameters. If experimental or observational data is used to develop fragility functions then correlations within the different parameters of the fragility functions could be due to similarities in the loading protocol, as well as uncertainties in the material properties and geometry of the components. For example, if a particular concrete beam

component tested in quasi-static loading has an EDP larger than mean EDP for DS_1: light cracking; then it is likely that, when tested further, the same component will have an EDP larger than the mean EDP for DS_2; severe cracking; (due to, for example, the particular beam having a higher than expected concrete tensile strength). Thus one reason for correlations between the epistemic uncertainties in fragility function parameters is due to dependence in the data used to develop them (i.e. violation of *randomisation*). Correlations between fragility function parameters can also be manifested during the mathematical fitting of the fragility functions (e.g. the correlations between the mean and standard deviation of Example 1 presented later in the chapter).

As epistemic uncertainty is related to the lack of knowledge regarding the seismic performance of a particular component then generally the epistemic uncertainty in different types of components will be uncorrelated. An exception would be related to similarities in the loading apparatus or loading regime used to test different components.

When expert judgement (i.e. method E in Porter et al. [8]) is used to estimate the parameters of a fragility function, which will involve significant epistemic uncertainty, then the correlation between parameters must also be based on expert judgement. Estimation of correlations between parameters using judgement can be aided by the consideration of common factors which affect the considered parameters [e.g. 30].

11.7.2 Determining correlations between different damage states

Determination of the correlation of epistemic uncertainties can be either accomplished using simulation, expert judgement, or both. For example, if bootstrap sampling is used to account for *finite sample uncertainty* there will be N realisations of the parameter values from which conventional correlation analyses [22, pg. 312] can be used to determine the correlations between the epistemic uncertainty.

If epistemic uncertainty is determined using a large degree of expert judgement then, as mentioned previously it is likely that expert judgement will also have to be used for the correlation. In this case it is important to note from Equation (11-8) that the assumption of a perfect positive correlation, typically in uncertainty analysis producing worse case assumption, will in fact lead to a reduction in the uncertainty in the damage state probability.

11.8 Combining different epistemic uncertainties

Below two methods are presented by which epistemic uncertainties from various

sources (e.g. *finite sample, loading protocol, host-to-target* etc.) can be combined to produce the total epistemic uncertainty in the fragility function parameters and their correlations.

11.8.1 Analytic approach to epistemic uncertainty combination

Since for finite sample uncertainty, the Student's-*t* and Chi-Square distributions tend to the normal distribution as the number of data increase, and purely for mathematical tractability with respect to other epistemic uncertainties it is advantageous to assume that the uncertainties in the mean and the variance of the fragility function are normally distributed. Based on this assumption, the uncertain value of the mean, $\mu_{\ln EDP|DS_i}$, is given by:

$$\mu_{\ln EDP|DS_i} = \sum_j \varepsilon_{\mu,j} + \overline{\mu_{\ln EDP|DS_i}} \tag{11-15}$$

where $\overline{\mu_{\ln EDP|DS_i}}$ is the initial estimate of the mean of the fragility function (based on, for example, data alone); and $\varepsilon_{\mu,j}$ represents the epistemic uncertainty in the mean, $\mu_{\ln EDP|DS_i}$, due to source *j*. The expectation and variance of the mean value of the fragility function are therefore:

$$\begin{aligned} E\left[\mu_{\ln EDP|DS_i}\right] &= \sum_j E\left[\varepsilon_{\mu,j}\right] + \overline{\mu_{\ln EDP|DS_i}} \\ Var\left[\mu_{\ln EDP|DS_i}\right] &= \sum_j Var\left[\varepsilon_{\mu,j}\right] \end{aligned} \tag{11-16}$$

where in the variance computation of Equation (11-16) the different sources of epistemic uncertainty are considered to be uncorrelated. The approach for the epistemic uncertainty in the variance, $\sigma^2_{\ln EDP|DS_i}$, is the same as that for the mean, and is not shown explicitly here. $\varepsilon_{\mu,j}$ will be a normal random variable with zero-mean (i.e. $E\left[\varepsilon_{\mu,j}\right]=0$) unless the particular source of epistemic uncertainty means that $\overline{\mu_{\ln EDP|DS_i}}$ gives a biased estimate of $\mu_{\ln EDP|DS_i}$. From Equation (11-15), the covariance between parameters *X* and *Y* (which may be the mean or the variance) for fragility functions *i* and *j*, respectively is:

$$\begin{aligned} Cov\left[X_{\ln EDP|DS_i}, Y_{\ln EDP|DS_j}\right] &= \sum_k \sum_l \rho_{\varepsilon_{X,k},\varepsilon_{Y,l}} \sigma_{\varepsilon_{X,k}} \sigma_{\varepsilon_{Y,l}} \\ &= \sum_k \rho_{\varepsilon_{X,k},\varepsilon_{Y,k}} \sigma_{\varepsilon_{X,k}} \sigma_{\varepsilon_{Y,k}} \end{aligned} \tag{11-17}$$

where the first line of Equation (11-17) comes from the definition of the covariance, while the second comes from the assumption that $\varepsilon_{X,k}$ and $\varepsilon_{Y,l}$ are uncorrelated if $k \neq l$. In the case of

$X = Y$ and $i = j$, Equation (11-17) reduces to the variance computation given in Equation (11-16).

11.8.2 Simulation approach to epistemic uncertainty combination

If, for example, bootstrap sampling is to be used to estimate *finite sample uncertainty* in a fragility function then it is relatively simple to also include other uncertainties into the computation simultaneously. As all the uncertainties would be considered simultaneously then the computation will yield the total epistemic uncertainties (i.e. mean, variance and covariance) in the parameters and hence there is no need combine the multiple sources of uncertainties as presented in the analytical method above. Unlike the analytic procedure where each source of epistemic uncertainty is described by using a random variable on either the mean or variance of the fragility function (e.g. $\varepsilon_{\mu,j}$ for epistemic uncertainty in the mean due to uncertainty source j), in the simulation approach, the uncertainties can be placed on individual data points. For example, the standard bootstrap sampling method creates a sample of N values from a dataset containing M values. In the case of multiple epistemic uncertainties, each of the M values in the dataset will be random variables with distributions corresponding to those defined for each type of uncertainty considered. Clearly, this simulation approach offers greater flexibility than the analytical approach above, but at the cost of additional complexity. For example, as the simulation is based on individual data and the epistemic uncertainties specified for each point, then the correlations between the uncertainties must also be specified at a point-by-point level. The trade-off between flexibility and simplicity will mean that both analytical and simulation-based approaches will be appropriate in different situations.

11.9 Fragility function and damage state uncertainty examples

In this section two examples are used to illustrate the aforementioned salient points regarding epistemic uncertainties in fragility functions and the resulting damage states probabilities. It is worth noting that the datasets of the examples presented are likely to be more substantial than those of many other components in practice, which should be borne in mind by the reader when examining the significance of the subsequent results.

11.9.1 Example 1: Fragility of hydraulic elevators

Porter [5] provides a fragility function for damage to hydraulic elevators based on observations from the Loma Prieta and Northridge earthquakes. Data from 91 elevators, including 10 that sustained damage was used to construct a lognormal distribution for the fragility of the elevators as a function of peak ground acceleration (PGA), which is shown in Figure 11-2. Because the observation data is of the *Bounding EDP* type as described previously, logistic regression was used to determine the values of the mean and standard deviation of the lognormal distribution.

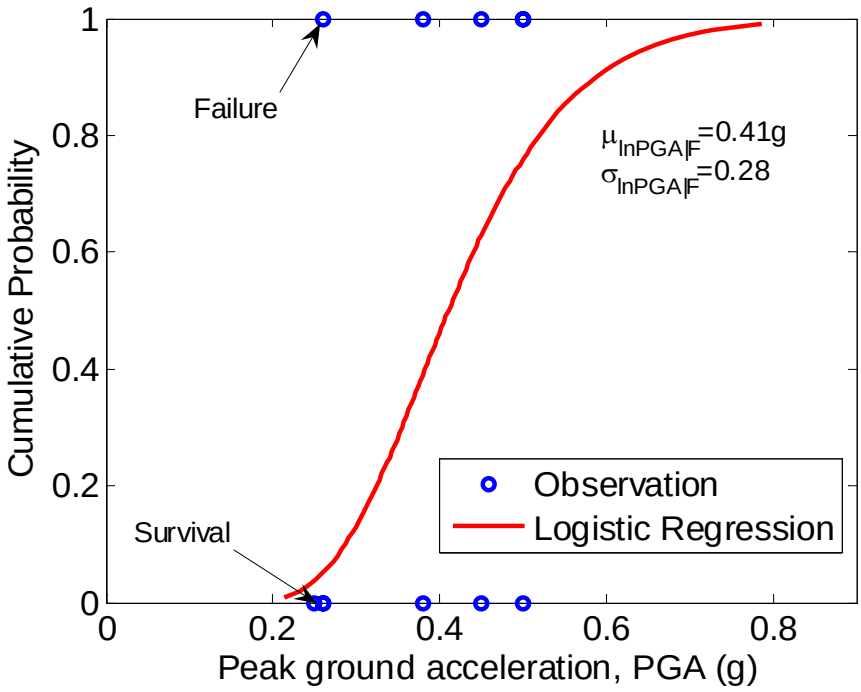


Figure 11-2: Comparison of the performance of Elevators during observed earthquakes and the fragility function obtained from logistic regression.

Using bootstrap sampling the epistemic uncertainty in the fragility function due to finite sample uncertainty can be computed. Figure 11-3a and Figure 11-3b illustrate histograms of the median and dispersion values obtained for each of the 500 bootstrap simulations conducted. It can be seen that the expected values of the median and dispersion of 0.418g and 0.284 are very similar to the values of 0.41g and 0.28 reported in Porter [5], indicating that there is little bias in the analytically-estimated mean value of these statistical moments. The uncertainty in the median and dispersion values, quantified using the coefficient of variation, were 0.21 and 0.42, respectively. In addition to the marginal distributions of the median and dispersion shown in Figure 11-3a and Figure 11-3b, Figure 11-3c illustrates that there is a significant correlation between the uncertainty in the median and dispersion. Figure 11-3d

illustrates the fragility functions that were obtained for each of the individual bootstrap simulations, as well as the mean, 16th and 84th percentiles of the cumulative probability for a given PGA value. It is firstly noted that the uncertainty in the fragility function is significantly less at low probabilities compared to that at high probabilities. This is due to the fact that of the 91 observations, 81 were 'no damage' observations, and hence the fragility function is relatively well constrained by the data at low cumulative probabilities. Conversely, the fact that there are only 10 'damage' observations is the cause of the larger uncertainty in the fragility function for large cumulative probabilities.

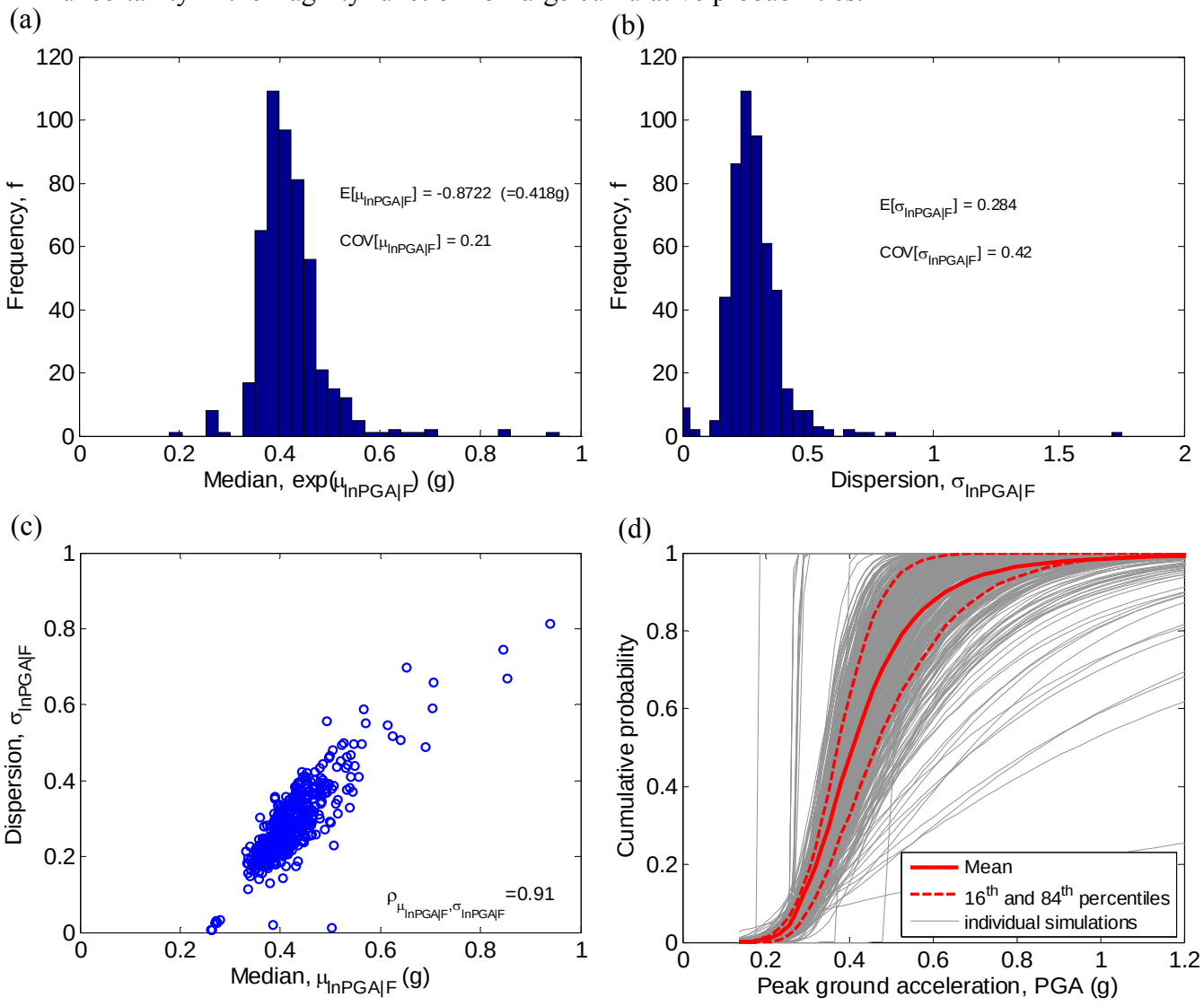


Figure 11-3: Bootstrap simulations of the empirical data used to create the Elevator fragility function in Porter [5]: (a) histogram of median; (b) histogram of dispersion; (c) correlation between median and dispersion; and (d) individual and mean ± one standard deviation fragilities.

In the previous sections, two options for the consideration of epistemic uncertainties were discussed. The first was where aleatoric and epistemic uncertainty are combined and propagated as a composite uncertainty (e.g. Equation (11-10)); while the second was where

epistemic and aleatory uncertainties are treated and propagated separately. Figure 11-4 compares the resulting mean fragility function obtained from the bootstrap simulation with the fragility function using the composite form of the aleatory and epistemic uncertainty. While both approaches match well over the central portion of the distribution, it can be seen that the composite fragility function over-estimates the probability of damage at the upper tail of the fragility distribution (i.e. the uncertainty based on the composite model is too low in this region). The reason for the error in the composite form of the fragility is that it is based on uncertainty only in the median of the distribution, however it is evident from Figure 11-1 that there is also significant uncertainty in the dispersion of the fragility and furthermore there is a strong correlation between these two moments.

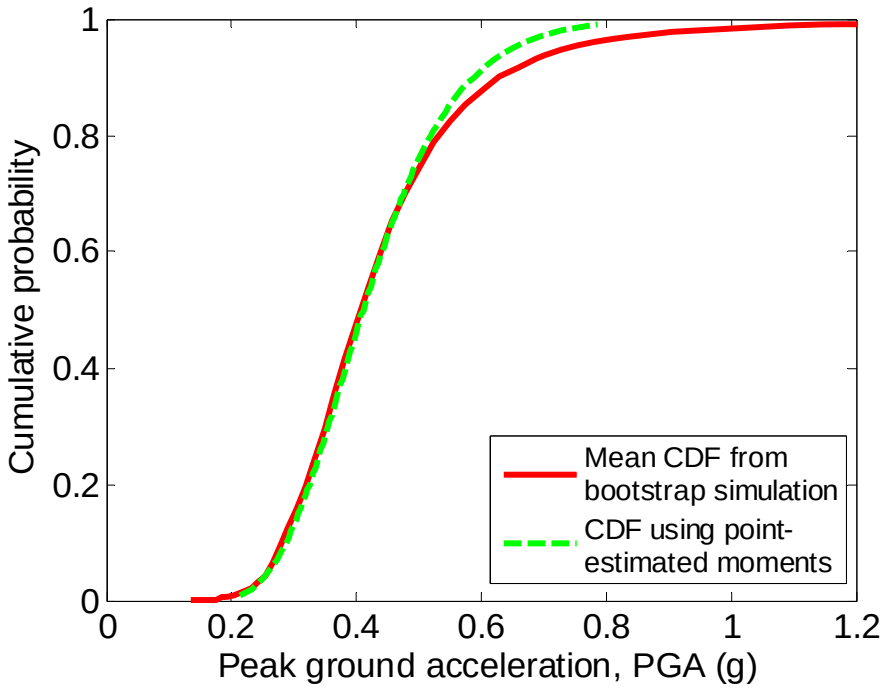


Figure 11-4: Comparison of the mean fragility from bootstrap simulation with that using only aleatory uncertainty and that with an SRSS combination of aleatory and epistemic uncertainty.

Figure 11-5 illustrates the second approach of separate consideration and propagation of epistemic uncertainties using bootstrap sampling (i.e. Figure 11-3d) and that obtained by using Monte Carlo (MC) simulation of the joint distribution of the median and variance based on Figure 11-3a-Figure 11-3c (i.e. that which could be used directly in a risk assessment). The distribution of the mean, $\mu_{\ln EDP|DS_i}$, and variance, $\sigma^2_{\ln EDP|DS_i}$, were simulated using a bi-variate normal distribution. That is, while Figure 11-3a-Figure 11-3c illustrate the median and dispersion (due to their familiarity with readers), it is the mean and variance of the distribution which are simulated. The reasons being the analytical basis previously discussed (i.e. that

relating to Equations (11-11) and (11-12)), and the fact that both the Student's-*t* and Chi-Square distributions asymptotically approach the normal distribution as the number of data increases. Thus, it should be noted that the correlation coefficient between $\mu_{\ln EDP|DS_i}$ and $\sigma^2_{\ln EDP|DS_i}$ used in the simulation was 0.80. Figure 11-5 illustrates that the bootstrap and MC simulation provide almost identical results in terms of the mean, 16th and 84th percentiles of the fragility function. It should be noted that the correlation between the median and dispersion is critical for the variation in the epistemic uncertainty with cumulative probability, as discussed above.

In regard to other sources of epistemic uncertainty, the nature of the data used (i.e. post-earthquake observations) ground motion variability (i.e. at the different elevator locations) and uncertainty in the in-situ construction are directly accounted for. However, the fact that 77 of the 91 data were obtained from Stanford University during the Loma Prieta earthquake [5] does mean that there is a strong dependence in the ground motion which these elevators were subjected to. Such correlations can be considered using methodologies such as those given in Straub and Der Kiureghian [29].

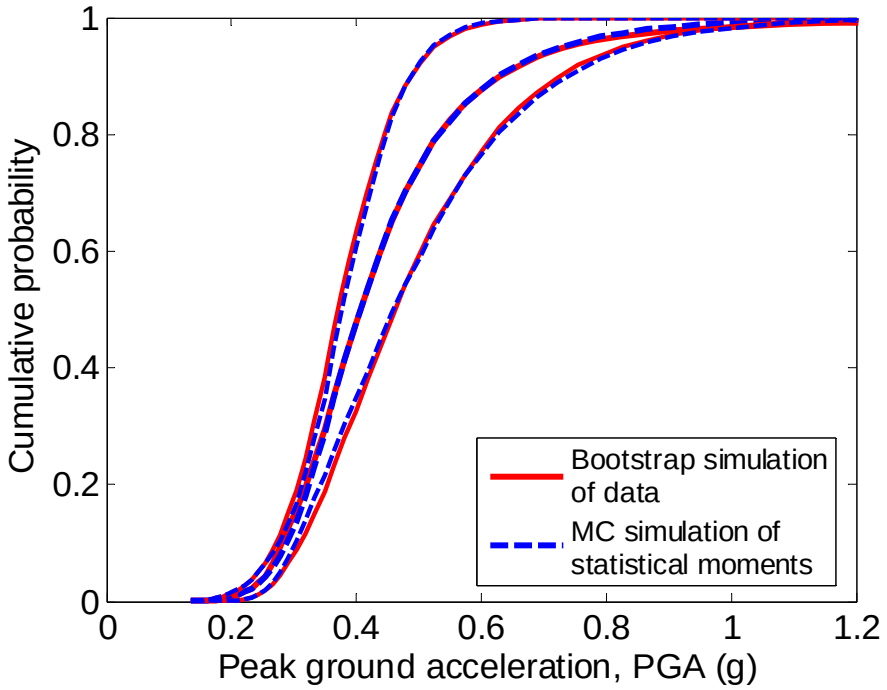


Figure 11-5: Comparison of the mean and ± one standard deviation values of the fragility function computing from bootstrap simulations and computed based on Monte Carlo (MC) simulation of the median and dispersion values.

11.9.2 Example 2: Fragility of slab column connections

Aslani and Miranda [2] provide fragility functions for damage to non-ductile slab-column connections based on quasi-static experimental tests performed over the last four decades. The damage states are defined as: (i) light cracking; (ii) severe cracking; (iii) punching shear failure; and (iv) loss of vertical carrying capacity (LVCC). A total of 82 experimental results were used, although not all tests were able to provide data for all of the four damage states. Table 11-2 provides the mean and variance of the parameters of the fragility functions for the first two damage states of the slab-column connections due to finite sample uncertainty and drift increment uncertainty in the inspection of damage. These two epistemic uncertainties were directly considered by Aslani and Miranda [2] using confidence intervals on the parameters. Figure 11-6 illustrates the mean and mean ± one standard deviation fragility functions for the two damage states based on bootstrap simulation (for finite sample uncertainty) and random residual errors for the drift increment uncertainty. In addition to the marginal distributions of the fragility function parameters, the correlation between the uncertainties in the parameters was given by:

$$\rho_{i,j} = \begin{array}{c} \\ \mu_1 \\ \sigma_1^2 \\ \mu_2 \\ \sigma_2^2 \end{array} \begin{array}{c} \begin{array}{cccc} \mu_1 & \sigma_1^2 & \mu_2 & \sigma_2^2 \end{array} \\ \begin{bmatrix} 1.00 & -0.04 & 0.16 & -0.06 \\ & 1.00 & 0.42 & 0.32 \\ & \textit{symmetric} & 1.00 & 0.08 \\ & & & 1.00 \end{bmatrix} \end{array} \tag{11-18}$$

In comparison to the correlation of 0.8 observed between the parameters for the hydraulic elevator fragility in the previous example, the correlations observed between the parameters of the first two DS's of the slab-column connections are somewhat smaller, with only the correlation between σ_1^2 and μ_2 ($\rho = 0.42$), and σ_1^2 and σ_2^2 ($\rho = 0.32$) being significant. This is likely due to the different methods that were used to develop the fragility functions: *bounding EDP* (method B) for the hydraulic elevators and *actual EDP* (method A) for the slab-column connections.

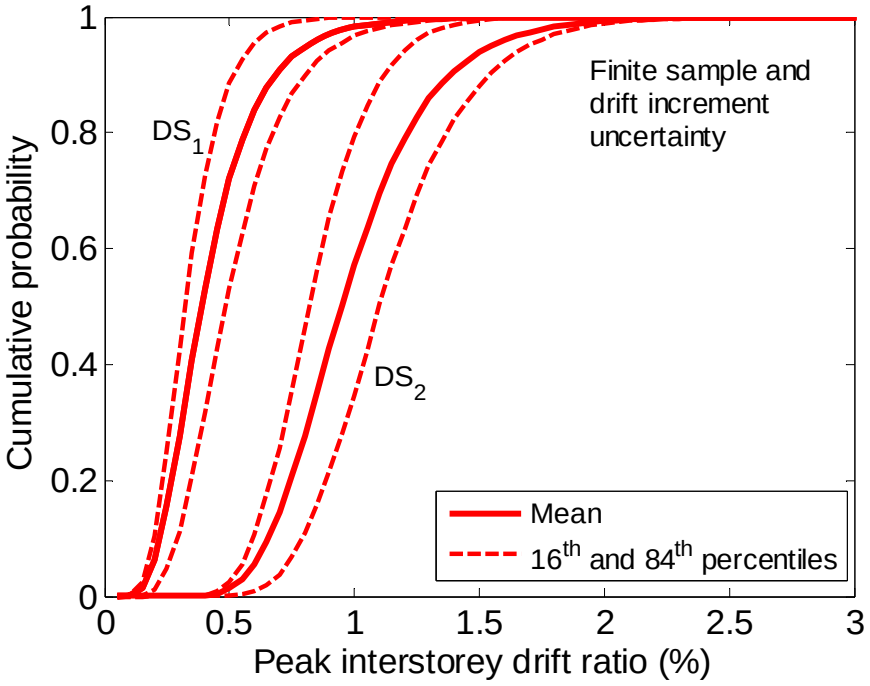


Figure 11-6: Mean and ± one standard deviation values due to finite sample and drift increment uncertainty for the first two fragility functions of slab-column connections [2].

Table 11-2: Epistemic uncertainty of parameters of slab-column fragility functions [2].

	N_{data}	$E\left[\mu_{\ln EDP\|DS_i}\right]$*	$Var\left[\mu_{\ln EDP\|DS_i}\right]$**	$E\left[\sigma^2_{\ln EDP\|DS_i}\right]$**	$Var\left[\sigma^2_{\ln EDP\|DS_i}\right]$**
DS_1	43	-1.11 (0.33)	0.123 (0.35)	0.152 (0.39)	$6.4\text{x}10^{-4}$ (0.025)
DS_2	33	-0.105 (0.90)	0.063 (0.25)	0.063 (0.25)	$2.9\text{x}10^{-4}$ (0.017)

*Values in brackets correspond to the median, $\overline{EDP|DS_i}$ (i.e. the exponent of the mean of $lnEDP|DS_i$)
**Values in brackets correspond to the standard deviation.

A potentially significant source of uncertainty not accounted for in Figure 11-6 is that of strain rate effects [25], given that the data on the damage states were obtained from quasi-static testing [2]. Quantification of the significance of such strain rate effects should be considered in practical applications, but is beyond the scope of this study.

Figure 11-7a illustrates the probability of being in damage state 1 as a function of interstorey drift ratio, based on the bootstrap simulations used to develop the fragility functions shown in Figure 11-6. In addition to the individual simulations, the mean, 16th and 84th percentiles are also depicted. Despite the significant number of specimens used to compute the fragility functions (i.e. 43 and 33 for DS_1 and DS_2, respectively), Figure 11-7a illustrates that there is still significant uncertainty in the probability being in DS_1 for a given

level of EDP. For example, at a peak interstorey drift ratio of 1% the mean probability is 0.42 with a 68% confidence interval (i.e. 16^{th} – 84^{th} percentiles) of 0.2-0.65. At a peak interstorey drift of 0.63%, where the probability of DS_1 is largest, the mean probability is 0.8 with a 68% confidence interval of 0.68-0.93. Figure 11-7b compares the mean, 16^{th} and 84^{th} percentiles obtained using MC simulation for the uncertainty in the fragility function parameters given in Table 11-2. The MC simulation was performed using both the correlation structure given by Equation (11-18) and also assuming uncorrelated uncertainties between the different distribution parameters. It can be seen that the assumption of no correlations is quite reasonable with only some disparity of the two results for peak interstorey drifts greater than 1%. This result is in contrast to that for the hydraulic elevators in the previous example were the correlation between the mean and standard deviation was significant. This illustrates that the significance of such correlations should be considered on a case-by-case basis.

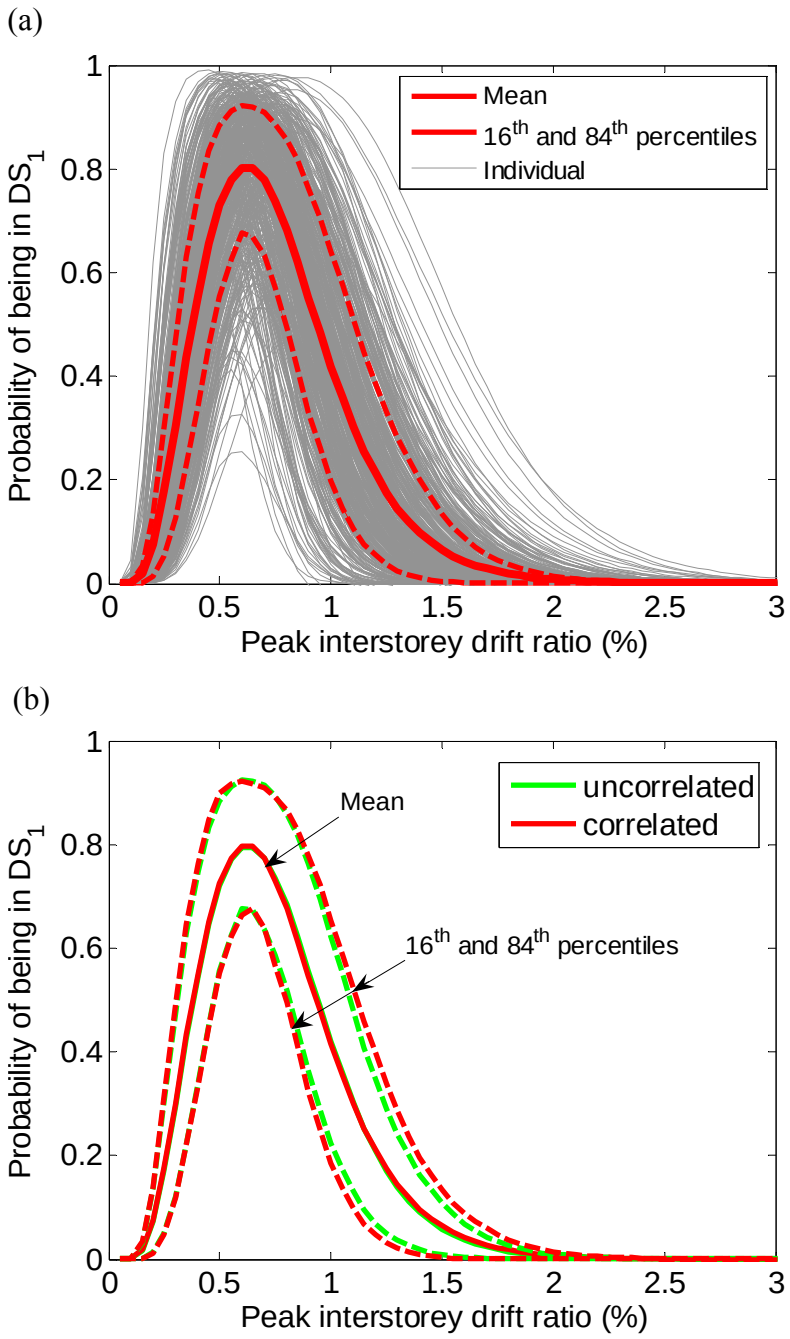


Figure 11-7: Probability being in DS_2 as a function of interstorey drift ratio: (a) effect of epistemic uncertainties; (b) effect of correlation between epistemic uncertainties on the probability obtained from MC simulation.

As previously mentioned, Aslani and Miranda [2] also developed fragility functions for the third and fourth damage states of punching shear failure and LVCC, respectively. In particular, for the punching shear failure damage state the use of the ratio of the gravity shear to the shear capacity, V_g/V_0, was used as an explanatory variable to reduce the uncertainty in the peak interstorey drift at the occurrence of punching shear failure. Figure 11-8 illustrates the relationship between interstorey drift at the occurrence of punching shear failure and the V_g/V_0 ratio for the 82 experimental tests used by Aslani and Miranda [2]. As an indication of

the significance of V_g/V_0, the dispersion in the fragility function without using V_g/V_0 was 0.62 [2], and the homoskedastic dispersion in Figure 11-8 is 0.31. Figure 11-8 also illustrates those data used in the fragility function development of punching shear failure which were from experiments performed by Hawkins *et al*. [31]. These 30 tests of Hawkins *et al*. [31] were monotonic quasi-static tests and therefore do not account for cyclic effects which will occur during ground motion shaking [27]. The dominance of the Hawkins *et al*. [31] data for V_g/V_0 ratios above 0.4 would mean that their removal would severely deplete the experimental data in this region (giving a large finite sample epistemic uncertainty). An alternative is the inclusion of bias and uncertainty factors to convert the monotonic results into equivalent 'cyclic' results similar to the remainder of the experimental data used. While the numerical values of such conversion factors are likely to be component-specific and are beyond the scope of this chapter, such a conversion will contain significant epistemic uncertainty. This epistemic uncertainty can be explicitly accounted for using the framework presented in this chapter.

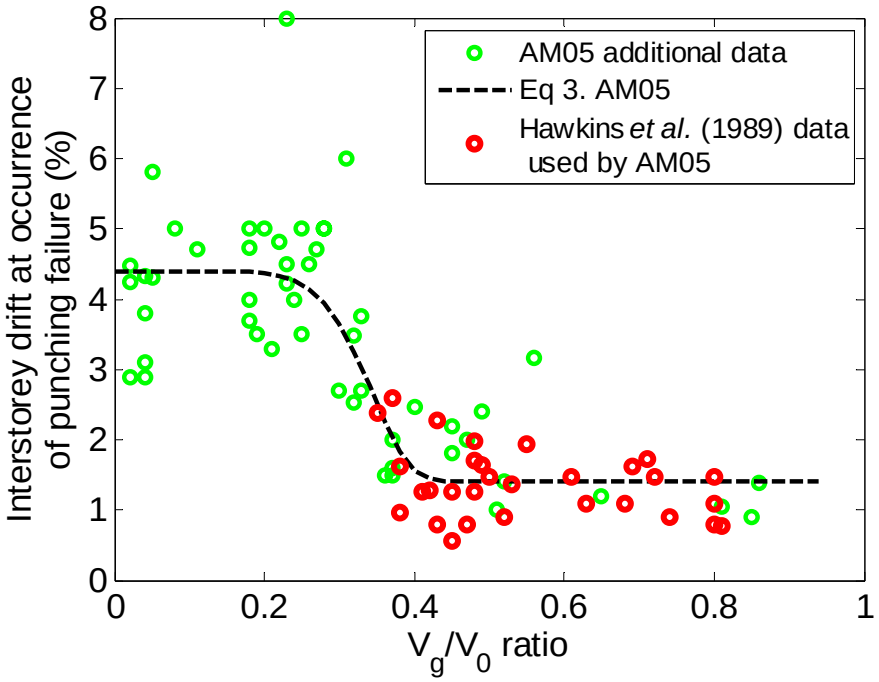


Figure 11-8: Dominance of the Hawkins *et al*. [31] data in the data set used by Aslani and Miranda [2] (AM05) at large V_g/V_0 ratios used for computing the punching shear failure fragility in slab-column connections. The Hawkins *et al*. [31] data is based on monotonic testing and therefore not directly representative of seismic demands due to earthquake induced shaking.

11.10 Fragility function quality

Table 6 of Porter *et al*. [8] provides various criteria by which one may assign a *quality* to the developed fragility function. From the framework, discussions and examples presented in this chapter it should be clear that the higher the *quality* of the fragility function, the smaller the epistemic uncertainty in the distribution parameters will be. The magnitude of epistemic uncertainties therefore acts as a quantitative measure of the *quality* of the fragility function. In addition, and more importantly, the benefit of reducing this epistemic uncertainty in performance assessments will clearly provide the need to improve poor *quality* fragility functions if they are critical to the seismic performance of the facility.

11.11 Conclusions

Fragility functions are an important step in the definition of the seismic performance of components needed in emerging performance-based earthquake engineering. Currently fragility functions are developed primarily using solely experimental, numerical or observational data, and therefore only capture aleatory uncertainty.

In this chapter various epistemic uncertainties in the development of component fragility functions have been identified. Methods for the consideration, combination and propagation of such epistemic uncertainties were discussed. The epistemic uncertainty associated with using finite datasets was illustrated via application to two components whose fragility functions have been published in technical literature.

There are clear benefits of the consideration of epistemic uncertainties which relate to the documentation, quality, implementation and updating of fragility functions and therefore such uncertainties should be addressed in future fragility functions developed for use in seismic performance assessment.

11.12 Acknowledgements

Financial support of the author from the New Zealand Tertiary Education Commission Bright Futures scheme, data provided by Dr. Hessam Aslani, and constructive comments by three anonymous reviewers are greatly appreciated.

11.13 References

[1] Ellingwood BR, Celik OC, and Kinali K. Fragility assessment of building structural systems in Mid-America. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1935-1952.

[2] Aslani H and Miranda E. Fragility assessment of slab-column connections in existing non-ductile reinforced concrete buildings. Journal of Earthquake Engineering 2005; 9(6): 777–804.

[3] Brown PC and Lowes LN. Fragility Functions for Modern Reinforced-Concrete Beam-Column Joints. Earthquake Spectra 2007; 23(2): 263-289.

[4] Badillo-Almaraz H, Wittaker AS, and Reinhorn AM. Seismic Fragility of Syspended Ceiling Systems. Earthquake Spectra 2007; 23(1): 21-40.

[5] Porter K. Fragility of Hydraulic Elevators for Use in Performance-Based Earthquake Engineering. Earthquake Spectra 2007; 23(2): 459-469.

[6] Hutchinson TC and Chaudhurib SR. Simplified expression for seismic fragility estimation of sliding-dominated equipment and contents. Earthquake Spectra 2006; 22(3): 709-732.

[7] Porter KA and Kiremidjian AS. Assembly-based vulnerability and its uses in seismic performance evaluation and risk-management decision-making. John A.Blume Earthquake Engineering Center, Stanford, CA, 2001. 216.

[8] Porter K, Kennedy R, and Bachman R. Creating Fragility Functions for Performance-based Earthquake Engineering. Earthquake Spectra 2007; 23(2): 471-489.

[9] ATC-58. Guidelines for seismic performance assessment of buildings: ATC-58 35% Draft. Applied Technology Council, 2007. http://www.atcouncil.org/atc-58.shtml (last accessed 19/11/08)

[10] FEMA-461. Interim Testing Protocols for Determining the Seismic Performance Characteristics of Structural and Nonstructural Components. Federal Emergency Management Agency, Washington, DC, 2007.

[11] Strasser FO, Abrahamson NA, and Bommer JJ. Sigma: Issues, insights, and challenges. Seismological Research Letters 2009; 80(1): 40-56.

[12] Der Kiureghian A and Ditlevsen O. Aleatory or epistemic? Does it matter? Structural Safety 2008; 31(2): 105-112, DOI: 10.1016/j.strusafe.2008.06.020.

[13] Der Kiureghian A. Non-erogodicity and PEER's framework formula. Earthquake Engineering and Structural Dynamics 2005; 34(13): 1643-1652, DOI: 10.1002/eqe.504.

[14] Kennedy RP, Cornell CA, Campbell RD, Kaplan S, and Perla HF. Probabilistic seismic safety study of an existing nuclear power plant. Nuclear Engineering and Design 1980; 59(2): 315-338.

[15] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[16] Kutner M, Nachtsheim C, Neter J, and Li W. Applied Linear Statistical Models. Fifth Edition ed. McGraw-Hill/Irwin: New York, 2005; 1396.

[17] Lilliefors H. On the Kolmorogov-Smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association 1967; 62: 399-402.

[18] Anagnos T. Development of an electrical substation equipment performance database for evaluation of equipment fragilities. Pacific Earthquake Engineering Center, 2001. 67. http://www.engr.sjsu.edu/tanagnos/Substation/index.htm

[19] Kythe PK and Schaferhotter MR. Handbook of Computational Methods for Integration. Chapman and Hall/CRC, 2000;

[20] Dolsek M. Incremental dynamic analysis with consideration of modelling uncertainties. Earthquake Engineering and Structural Dynamics 2009, DOI: 10.1002/eqe.869.

[21] Cornell CA, Jalayer F, Hamburger RO, and Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. Journal of Structural Engineering 2002; 128(4): 526–533.

[22] Ang AHS and Tang WH. Probability concepts in engineering: Emphasis on applications in civil and environmental engineering. John Wiley & Sons, 2007;

[23] Wasserman L. All of statistics: a concise course on statistical inference. Springer: New York, 2004; 442.

[24] Leon RT and Deierlein GG. Considerations for the use of quasi-static testing. Earthquake Spectra 1996; 12(1): 87-109.

[25] Abrams D. Effects of scale and loading rate with tests of concrete and masonry structures. Earthquake Spectra 1996; 12(1): 13-28.

[26] Montgomery DC. Design and analysis of experiments. John Wiley & Sons: New York, 1991;

[27] Krawinkler H. Cyclic loading histories for seismic experimentation on structural components. Earthquake Spectra 1996; 12(1): 1-12.

[28] Williams MS and Sexsmith RG. Seismic damage indicies for concrete structures: A state-of-the-art review. Earthquake Spectra 1995; 11(2): 319-349.

[29] Straub D and Der Kiureghian A. Improved seismic fragility modelling from empirical data. Structural Safety 2008; 30(4): 320-336.

[30] Hudak D and Maxwell M. A macro approach to estimating correlated random variables in engineering production projects. Construction Management and Economics 2007; 25: 883-892.

[31] Hawkins NM, Bao A, and Yamazaki J. Moment transfer from column slabs to columns. ACI Structural Journal 1989; 86(6): 705-716.

12.Ground Motion Prediction Equation for Spectrum Intensity Based on Spectral Acceleration Equations

Bradley BA, Cubrinovski M, MacRae GA, Dhakal RP. Ground motion prediction equation for spectrum intensity based on spectral acceleration equations. *Bulletin of the Seismological Society of America* 2009; **99**(1): 277-285.

12.1 Abstract

Spectrum intensity (*SI*) defined as the integral of the pseudo spectral velocity of a ground motion from 0.1 to 2.5 seconds, has recently been shown to be an intensity measure that efficiently predicts the seismic response of both liquefiable and non-liquefiable soil deposits, as well as the seismic demands on pile foundations embedded in such deposits. In order for such an intensity measure to be used in performance-based assessment and design, ground motion prediction relations are required to develop ground motion hazard curves in terms of *SI* for various sites. As such relationships developed specifically for *SI* are sparse, the authors propose the development of a relationship based on current ground motion prediction relations for spectral acceleration, which are available in most regions of seismic activity. Comparison with a direct prediction equation for *SI* provides a validation of the proposed approach. It is illustrated that *SI* is an intensity measure with a good predictability, thereby further promoting its attractiveness for use in reliability-based seismic response analysis.

12.2 Introduction

The seismic demand on structures due to ground motion excitation is highly uncertain due to the inherently random nature of the fault rupture process, seismic wave propagation and local site effects, as well as the variation in the seismic response of structures subjected to ground motion excitations of similar intensity. The latter indicates that in such seismic response analysis it is important to employ a ground motion intensity measure (IM) which is *efficient* [1] in predicting these seismic demands. The former indicates however, that consideration should also be given to the *predictability* [2] of the IM, which relates to the ability of the IM to be predicted from available ground motion prediction equations. Such a ground motion prediction equation is required to determine the ground motion hazard curve for a specific site, allowing determination of the temporal occurrence of the adopted ground motion IM. In addition to having a ground motion prediction equation for a specified IM, it is also desirable that the prediction equation has a high predictability (i.e. a relatively small variation in the ground motion intensity for a given earthquake scenario) which influences the ground motion hazard curve determined via probabilistic seismic hazard analysis (PSHA). Kramer and Mitchell [2] illustrate how significant uncertainty in the ground motion prediction model of the IM causes an increase in the ground motion hazard and this in turn results in an increase in the demand hazard. In other words, if the *predictability* of an IM is poor, then the accuracy in predicting the seismic response (for a given earthquake scenario) will also be poor.

Recent research [3, 4] has suggested that Spectrum Intensity (*SI*) is an *efficient* predictor of the seismic response of soil deposits, both for liquefiable and non-liquefiable soils, as well as the seismic demands of pile foundations embedded in such soil deposits. As mentioned above, however, in order for *SI* to be routinely used in performance-based assessment and design, ground motion prediction relations are required to develop ground motion hazard curves in terms of *SI* at various sites of interest. Such relationships developed specifically for *SI* are sparse, thus limiting the potential application of *SI* in performance-based seismic assessments.

This chapter introduces an indirect method for development of a ground motion prediction equation for *SI* based on ground motion prediction equations for spectral accelerations which are available in abundance. The formulation presented herein allows computation of *SI* based on any spectral acceleration (*Sa*) prediction equation as well as to

incorporate specific features of individual *Sa* prediction models.

12.3 Spectrum intensity prediction equation

Spectrum intensity, *SI*, originally proposed by Housner [5, 6] is defined as the integral of the pseudo-spectral velocity (*PSV*) over the period range of 0.1-2.5 seconds as given by Equation (12-1). In general, *SI*, is defined for any level of viscous damping, however commonly a value of 5% of the critical damping is selected, and will be adopted here.

$$SI = \int_{0.1}^{2.5} PSV(\mathrm{T},5\%)\mathrm{dT} \tag{12-1}$$

The following section presents a method from which a ground motion prediction equation for *SI* can be determined from ground motion prediction equations for spectral acceleration, (*Sa*), which are readily available and well developed.

In general, ground motion prediction equations for spectral accelerations provide the median (50^{th} percentile) spectral acceleration and an associated lognormal standard deviation. The lognormal standard deviation is typically provided as it has been shown by various researchers that spectral accelerations are approximately lognormally distributed [7].

In order to compute the ground motion prediction equation for *SI* as proposed in this chapter, it will be necessary to make use of the non-log form for the (statistical) moments of the spectral acceleration ordinates. Thus, Equations (12-2) and (12-3), which are properties of the lognormal distribution, can be used to obtain the non-log moments of the spectral acceleration prediction equation:

$$\mu_{Sa} = Sa_{50} \exp\left(\frac{1}{2}\sigma^2_{\ln Sa}\right) \tag{12-2}$$

$$\sigma_{Sa} = \mu_{Sa}\sqrt{\exp\left(\sigma^2_{\ln Sa}\right)-1} \tag{12-3}$$

where Sa_{50} and $\sigma_{\ln Sa}$ are the median and lognormal standard deviation of the spectral acceleration, determined directly from *Sa* prediction equations; and μ_{Sa} and σ_{Sa} are the (non-log form) mean and standard deviation of the spectral acceleration, respectively.

Further noting that the pseudo-spectral velocity is related to the spectral acceleration by:

$$PSV(T_i) = \frac{Sa(T_i)}{\omega_i} = \frac{Sa(T_i).T_i}{2\pi} \tag{12-4}$$

then the mean and standard deviation of *PSV* for a given earthquake scenario (magnitude,

distance, fault style and site conditions) and vibration period can be obtained from:

$$\mu_{PSV_i} = \mu_{Sa_i} / \omega_i \tag{12-5}$$

$$\sigma^2_{PSV_i} = \sigma^2_{Sa_i} / \omega_i^2 \tag{12-6}$$

Replacing the integral of Equation (12-1) with a discrete summation approximation gives:

$$SI = \Delta T \sum_{i=1}^{n} w_i PSV(\mathrm{T_i}, 5\%) \tag{12-7}$$

where n = the number of periods (from 0.1 to 2.5 seconds) that PSV is computed at; ΔT is the size of the vibration period discretization (the step-size used in the integration); and w_i are integration weights which depend on the integration scheme used. From Equation(12-7), the mean and variance of SI can then be computed from [8]:

$$\mu_{SI} = \Delta T \sum_{i=1}^{n} w_i \mu_{PSV_i} \tag{12-8}$$

$$\sigma^2_{SI} = (\Delta T)^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \left(w_i w_j \rho_{PSV_i, PSV_j} \sigma_{PSV_i} \sigma_{PSV_j} \right) \tag{12-9}$$

where $PSV_i = PSV(\mathrm{T_i}, 5\%)$; μ_{PSV_i} and $\sigma^2_{PSV_i}$ are the mean and variance, respectively, of the PSV ground motion prediction equation; and ρ_{PSV_i, PSV_j} is the correlation between PSV_i and PSV_j, i.e. the correlation between the PSV at two different vibration periods. Note that Equations (12-8) and (12-9) are exact, irrespective of the distribution of the spectral velocity terms. However, Equations (12-8) and (12-9) provide only the first two (statistical) moments of SI and no information on the resulting distribution of SI.

Again it is noted that as $PSV_i = Sa_i/\omega_i$, where ω is the circular natural frequency, then it can be shown that (i.e. Appendix D):

$$\rho_{PSV_i, PSV_j} = \rho_{Sa_i, Sa_j} = \frac{\exp\left(\rho_{\ln Sa_i, \ln Sa_j} \sigma_{\ln Sa_i} \sigma_{\ln Sa_j}\right) - 1}{\sqrt{\exp\left(\sigma^2_{\ln Sa_i}\right) - 1} \sqrt{\exp\left(\sigma^2_{\ln Sa_j}\right) - 1}} \approx \rho_{\ln Sa_i, \ln Sa_j} \tag{12-10}$$

where ρ_{Sa_i, Sa_j} is the correlation between spectral accelerations at vibration periods i and j; and $\rho_{\ln Sa_i, \ln Sa_j}$ is the correlation between the logarithm of spectral accelerations at vibration periods i and j. Equation (12-10) illustrates that the correlation between two different spectral velocities is equivalent to the correlation between two different spectral acceleration terms, and is a first order approximation to the correlation between the logarithm of two different

spectral acceleration terms (although note that the exact expression for the relationship between log and non-log correlations is used herein unless otherwise stated).

Numerous spectral acceleration relationships can therefore be used to compute the mean and standard deviation of *Sa* (which can then be converted to *PSV*), while several models are available for $\rho_{\ln Sa_i, \ln Sa_j}$ [9-11].

Equations (12-8) and (12-9) gives the two (non-log) moments for *SI* and if it is assumed that the distribution of *SI* can be adequately represented by the lognormal distribution (which is shown to be the case later in this chapter) then the median and lognormal standard deviation can be computed from re-arranged forms of Equations (12-2) and (12-3):

$$SI_{50} = \frac{\mu_{SI}^2}{\sqrt{\sigma_{SI}^2 + \mu_{SI}^2}} \tag{12-11}$$

$$\sigma_{\ln SI} = \sqrt{\ln\left(\left(\frac{\sigma_{SI}}{\mu_{SI}}\right)^2 + 1\right)} \tag{12-12}$$

where μ_{SI} and σ_{SI} are obtained from Equations (12-8) and (12-9), respectively, and SI_{50} and $\sigma_{\ln SI}$ are the median and lognormal standard deviation (dispersion) of the spectrum intensity of a ground motion produced for a given scenario. Equations (12-11) and (12-12) can be used directly in conventional *PSHA* computer programs. The principal benefit of computation of *SI* directly from spectral acceleration prediction equations comes from the significantly advanced state of *Sa* prediction equations in regard to quantification of faulting styles and site conditions, and large databases of empirical data used for their calibration [e.g. 12, 13, 14].

12.4 Distribution of *SI* prediction equation

The previous section illustrated how the (non-log form) mean and standard deviation of the spectrum intensity can be determined from spectral acceleration prediction equations. However, it was also mentioned that while the mean and standard deviation are correct using Equations (12-8) and (12-9), respectively, no information is given regarding the distribution of *SI*. In what follows, the distribution of *SI* using the proposed approach presented in the previous section is investigated.

In order to determine the distribution shape of *SI*, a Monte-carlo scheme is used to: (i) randomly generate correlated logarithmic spectral acceleration amplitudes at various periods; (ii) take the exponent of the logarithmic terms to obtain (non-log) *Sa* terms; (iii) convert to

pseudo-spectral velocities using Equation (12-4); and (iv) compute the spectrum intensity, *SI*. These four steps are further elaborated below.

12.4.1 Step 1: Generate correlated logarithmic spectral acceleration terms

As spectral accelerations are known to be correlated at various periods, then random generation of vectors of spectral acceleration terms involves generating correlated random variables. It has already been mentioned that logarithmic spectral acceleration ordinates have been shown by various researchers to be normally distributed. In addition, Jayaram and Baker [7] illustrate that logarithmic spectral acceleration terms are not only marginally, but also jointly normally distributed. Thus, it is possible to generate the logarithmic spectral acceleration terms using a multivariate normal distribution with correlation matrix as defined by an empirical equation for $\rho_{\ln Sa_i, \ln Sa_j}$. Herein, unless otherwise noted, the Boore and Atkinson [13] NGA ground motion prediction equation for spectral acceleration, and the Jayaram and Baker [7] correlation model are used, both of which were developed using the same ground motion database.

12.4.2 Step 2: Convert to non-log spectral acceleration

The correlated logarithmic spectral acceleration terms in the previous section can be simply converted to (non-log) spectral accelerations using:

$$Sa_i^k = \exp\left(\ln Sa_i^k\right) \tag{12-13}$$

where $\ln Sa_i^k = \ln Sa(T = T_i, 5\%)^k$ is the logarithmic spectral acceleration at vibration period T_i, generated during realisation k; and Sa_i^k is the non-log form of $\ln Sa(T = T_i, 5\%)^k$.

12.4.3 Steps 3&4: Convert to spectral velocity and compute SI_i.

Once the vector of Sa_i^k terms have been obtained, Equation (12-4) can be used to obtain a vector of pseudo-spectral velocity terms. Equation (12-7) can then be used to compute the spectrum intensity for realization k.

12.4.4 Resulting distribution

Figure 12-1 illustrates a comparison of the distribution of *SI* obtained using the simulation procedure described above with the lognormal and normal analytic distributions obtained based on the moments computed using Equations (12-8) and (12-9). The figure is based on a $M_w = 6.5$ strike-slip rupture at a distance of 30 km from a site with a 30m-weighted-average shear wave velocity of $V_{s(30)} = 300$ m/s. The simulation is based on 10000 realisations using the approach outlined in the previous section. In both the simulation and the analytic method (using Equations (12-8) and (12-9) for the normal distribution and Equations (12-11) and (12-12) for the lognormal distribution) a vibration period discretization of 0.1 seconds was adopted, (the vibration period discretization is discussed later in the chapter). It can be seen that the lognormal distribution provides an excellent approximation to the results of the simulation while the normal distribution provides a poor fit. It is worth noting here that while the central limit theorem [8] may intuitively suggest that as the vibration period discretization becomes significantly small (i.e. the number of terms in the summation in Equations (12-8) and (12-9) become significantly large) then the resulting distribution for *SI* would approach the normal distribution (in the case of $\Delta T = 0.1$, there are 25 terms in the summation). However, as illustrated by the results of the simulation the central limit theorem does not apply here. This is primarily due to the high linear dependence (correlation) between the spectral acceleration terms at different vibration periods. Although not shown here, it was verified by the authors that reducing the vibration period discretization as low as $\Delta T = 0.01$, such that there are 250 spectral acceleration terms in Equation (12-7), did not affect the resulting distribution of *SI*.

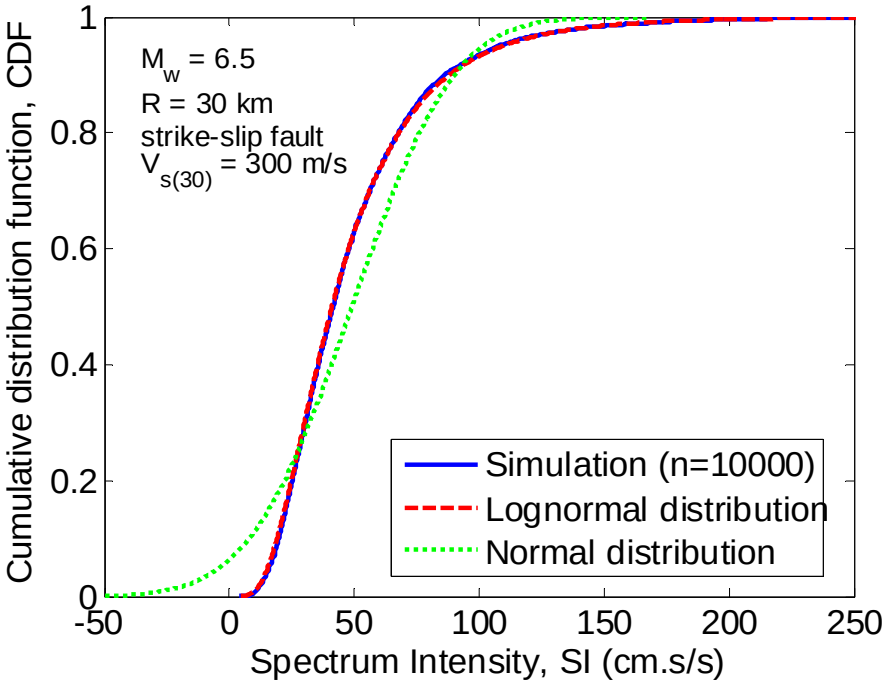


Figure 12-1: Distribution of *SI* obtained by simulation compared with normal and lognormal distributions.

As previously noted Equations (12-8) and (12-9) have been used to obtain the moments for *SI*, which have then been used to obtain the normal and lognormal distributions shown in Figure 12-1. Figure 12-2a illustrates the error in the prediction of the mean and standard deviation using Equations (12-8) and (12-9) relative to the results of the simulation. Because of the random nature of the simulations, multiple simulations (20 in this case) each with 10000 realizations were used to obtain cumulative distributions of the relative errors. It can be seen that the prediction of the mean *SI* from Equation (12-8) is accurate with the majority of the 20 different simulations giving an absolute relative error of less than 0.5%. The distribution of the relative error in the standard deviation is also centered around zero, but with a slightly larger scatter than that of the mean (if the first order approximation is made for the correlation (Equation (12-10)) there is a bias in the standard deviation of approximately 2.5%).

Figure 12-2b illustrates the variation in the Kolmogorov-Smirnov (KS) test statistic [8] over the 20 different simulations performed. The KS statistic gives the maximum difference between the empirical (from simulation) and analytical (lognormal) distributions over all the range of spectrum intensity values obtained. Thus Figure 12-2b indicates that over the entire range of *SI* values, the maximum difference between the empirical distribution and the lognormal distribution is typically about 0.011. The KS critical test statistic value for

n = 10000 simulations is also shown for a confidence level of α = 5%. The fact that 80% of the simulations performed fall within the KS critical value indicates further that while the lognormal distribution is not exact, the assumption is relatively good.

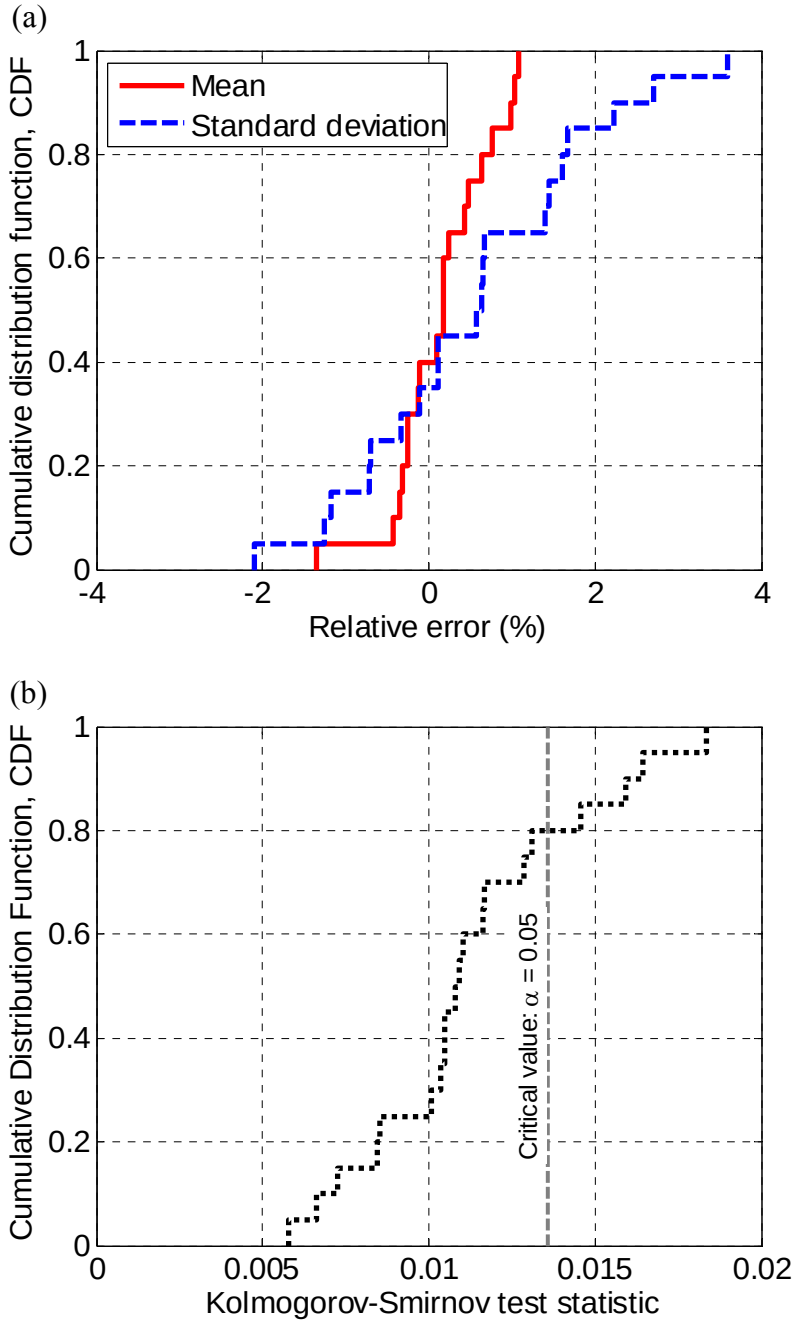


Figure 12-2: Errors in the SI distribution simulation: (a) relative error in the mean and standard deviation; and (b) the Kolmogorov-Smirnov test statistic.

Because the scale used in the KS test is arithmetic, it is unlikely that the value of the statistic is obtained near the tails of the distributions [8]. Thus, separate attention should be devoted to the inspection of the tails of the empirical distribution. Figure 12-3a illustrates the comparison between the tails of the empirical and lognormal distributions for the same

simulation as shown in Figure 12-1. It can be seen that the comparison between the empirical distribution from simulation and the lognormal distribution based on Equations (12-8) and (12-9) is good over the full tail of the *SI* distribution. The Anderson-Darling (AD) goodness-of-fit test [8] is one such test for testing the tails of empirical distributions. In similar fashion to the KS statistic computed above, 20 different simulations using 10000 realisations were used to compute the empirical distribution for the AD test statistic shown in Figure 12-3b. Also for comparison the critical AD test statistic for a confidence level of 5% is shown. The fact that approximately 90% of the 20 simulations gave an AD test statistic less than the critical indicates the adequacy of the lognormal distribution in describing the tails of the *SI* distribution.

This section has investigated the assumption of lognormality for the distribution of *SI* obtained based on Equations (12-8) and (12-9). It has been illustrated that both the 'body' and 'tail' of the distribution conform to the lognormal distribution based on statistical tests. While there is some small error between the empirical and lognormal distributions, it should be realised that the assumption that spectral accelerations are lognormally distributed, although shown to be valid, is still an assumption. Also, the error induced in assuming that *SI* is lognormally distributed is likely to be negligible compared to the variation in ground motion due to epistemic uncertainties associated with earthquake rupture forecasts (ERFs) and the functional forms for spectral acceleration attenuation relationships themselves.

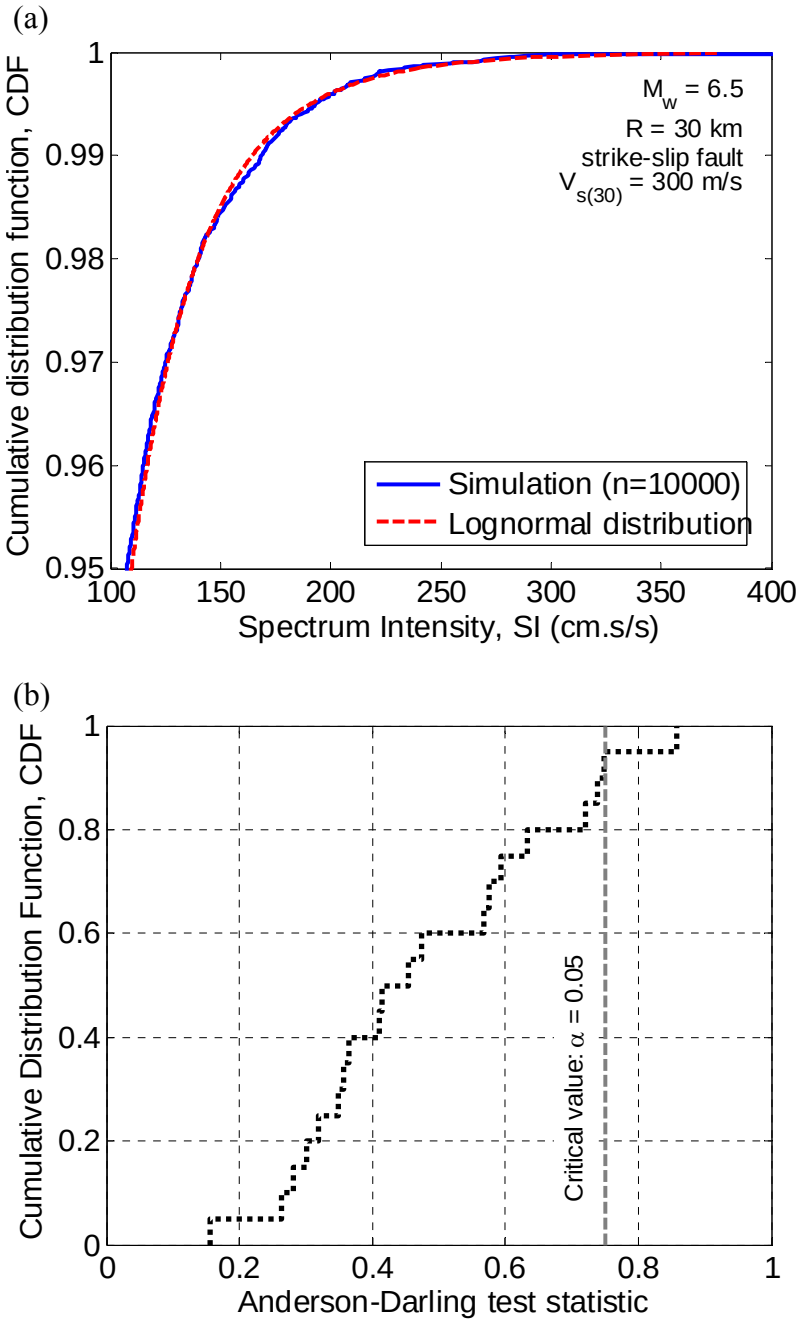


Figure 12-3: Investigation of the tail of the *SI* distribution: (a) comparison between empirical and lognormal distributions for the simulation shown in Figure 12-1; and (b) the Anderson-Darling test statistic.

12.5 Effect of vibration period discretization

The discrete form for the computation of *SI*, as given by Equation (12-7) requires the size of the vibration period discretization, ΔT, to be selected in order to (discretely) obtain *SI* from the integral of the pseudo-velocity spectra. The selection of the discretization size will have an effect on both the median and dispersion (lognormal standard deviation) obtained

using Equations (12-11) and (12-12) (as will the selection of the numerical integration method used, which affects the integration weights, w_i).

Since empirical ground motion prediction equations use regression to determine the distribution of spectral acceleration at a discrete range of periods, then interpolation is required to determine the distribution of spectral ordinates at periods other than those for which empirical coefficients are determined for. The simplest and most common form of interpolation in such cases is linear interpolation, and is therefore adopted here. Because linear interpolation is used to determine the distribution of spectral accelerations at various periods, it seems appropiate that the Trapezoidal rule is used to evaluate Equation (12-1), which also makes use of linear interpolation between integration points. Thus, using the Trapezoidal rule in the form of Equation (12-7), the integration weights, w_i, will take the value of 0.5 for i=1 and n, and 1.0 otherwise.

Figure 12-4 illustrates the convergence of the median and dispersion for *SI* using the proposed approach. A range of discretization sizes from 1.2s (giving 3 *Sa* points), through to 0.01s (giving 250 *Sa* points), were considered. Figure 12-4 illustrates that, as expected, convergence is achieved as the discretization size is reduced, and that a step size below ΔT = 0.2s is appropriate for a wide range of magnitude and distance scenarios (two of which are shown here).

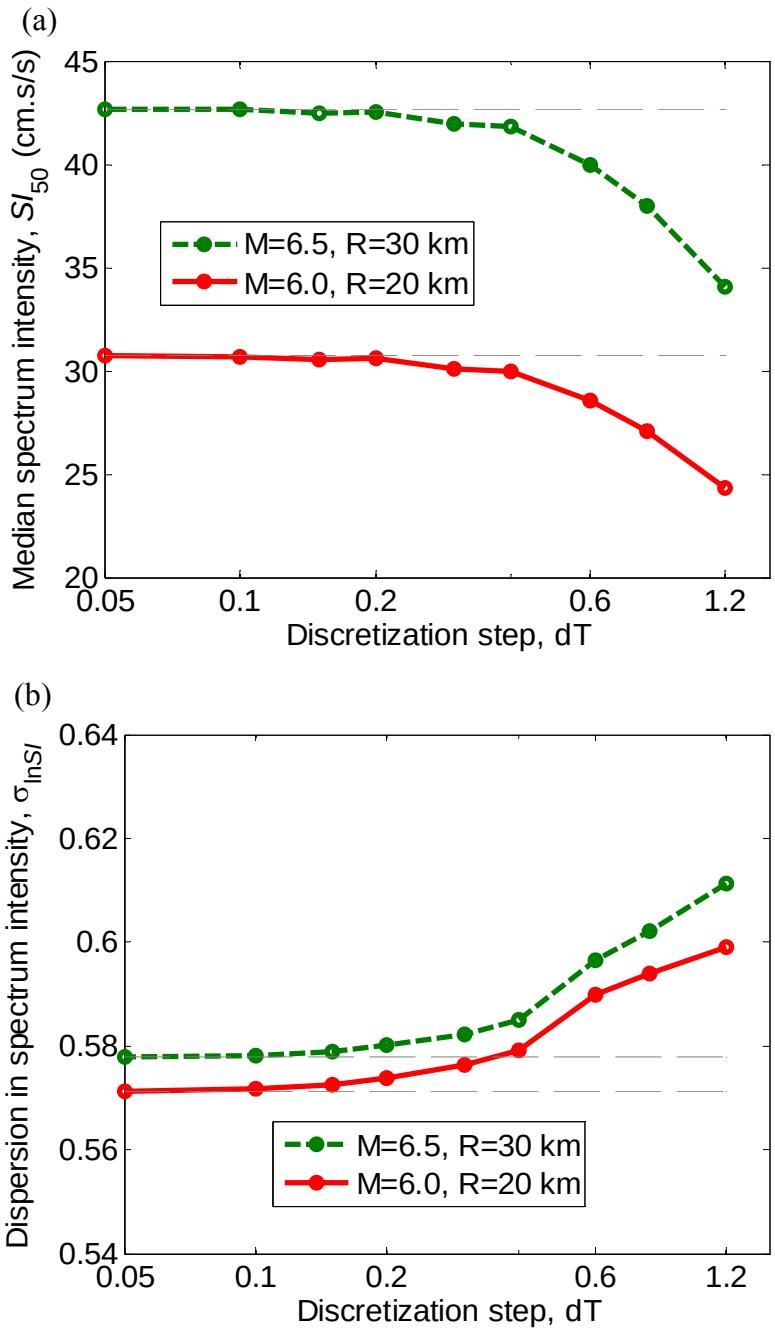


Figure 12-4: Effect of period discretization size on convergence of: (a) median (Equation (12-11)); and (b) lognormal standard deviation (dispersion) (Equation (12-12)).

12.6 Properties of *SI* prediction equation

12.6.1 Comparison with direct prediction of *SI*

It is of interest to compare the prediction of *SI* based on a prediction relationship for *SI*,

with the prediction of *SI* obtained via prediction equations for *Sa*, which is the focus of this study. Danciu and Tselentis [15] present ground motion prediction equations for various ground motion intensity measures for predominantly shallow earthquake ruptures in Greece of normal and thrust focal mechanisms. The ground motion intensity measures considered include both spectrum intensity, *SI*, and spectral acceleration, *Sa*. Danciu and Tselentis [15] use a slightly modified version of *SI* than that presented here, their version having a 'normalising' factor of '1/2.4' at the front of the integral given in Equation (12-1). It is a trivial matter to multiply the results obtained from Danciu and Tselentis [15] by 2.4 to obtain those results presented herein. In addition, the standard deviations given by Danciu and Tselentis [15] are in terms of the base 10 logarithm ($\log_{10}$), and are converted to natural logarithms (ln) using a scaling factor of ln(10). Figure 12-5 illustrates the comparison between the median, 16th and 84th percentiles of the spectrum intensity distribution for a $M_w = 6.5$ normal fault rupture at various distances from a rock site. It is evident from the figure that the median spectrum intensities obtained using the two different approaches are practically identical, while there is a minor under-prediction of the magnitude of the dispersion.

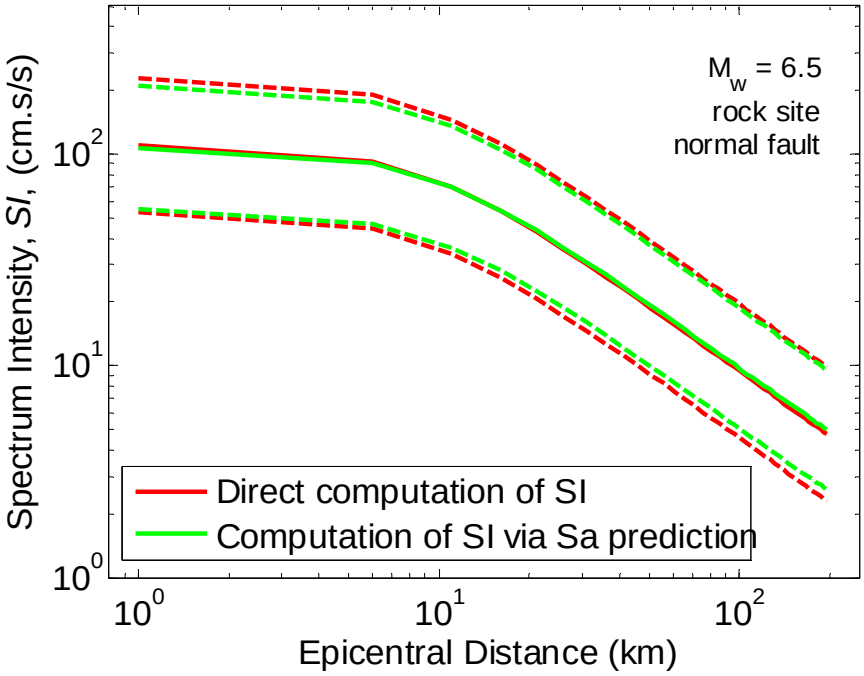


Figure 12-5: Comparison of *SI* obtained directly from an *SI* prediction equation and that obtained by the proposed method via spectral acceleration prediction equations, both from Danciu and Tselentis [15].

12.6.2 Size of lognormal standard deviation of *SI*

As was mentioned in the introduction of this chapter, it is desirable to have a ground motion intensity measure (IM) which is both *efficient* in predicting the response of a structure due to ground motion excitation, and *predictable* in the sense that a ground motion prediction equation exists which can estimate the IM with a low (relative to alternate IMs) uncertainty for a given earthquake scenario. Thus it is pertinent to investigate the predictability of *SI* in comparison to other common IMs such as peak ground velocity (*PGV*), and *Sa*. Figure 12-6 provides a comparison between the dispersions of the three aforementioned IMs using the Boore and Atkinson [13] NGA prediction equation for computation of *Sa* and *PGV*, and the proposed approach for computing *SI* with the Boore and Atkinson prediction equation for *Sa*. As is generally observed in spectral acceleration prediction equations the dispersion increases as the period of vibration increases, while *PGV* and *SI* are obviously independent of vibration period. Both *PGV* and *SI* are observed to be more *predictable* than the *Sa* terms. For *SI* in particular, this is due to the fact that spectral acceleration terms are not perfectly correlated at various vibration periods, which reduces that standard deviation in *SI* compared to the case of perfectly correlated *Sa* terms (e.g. Equation (12-12)).

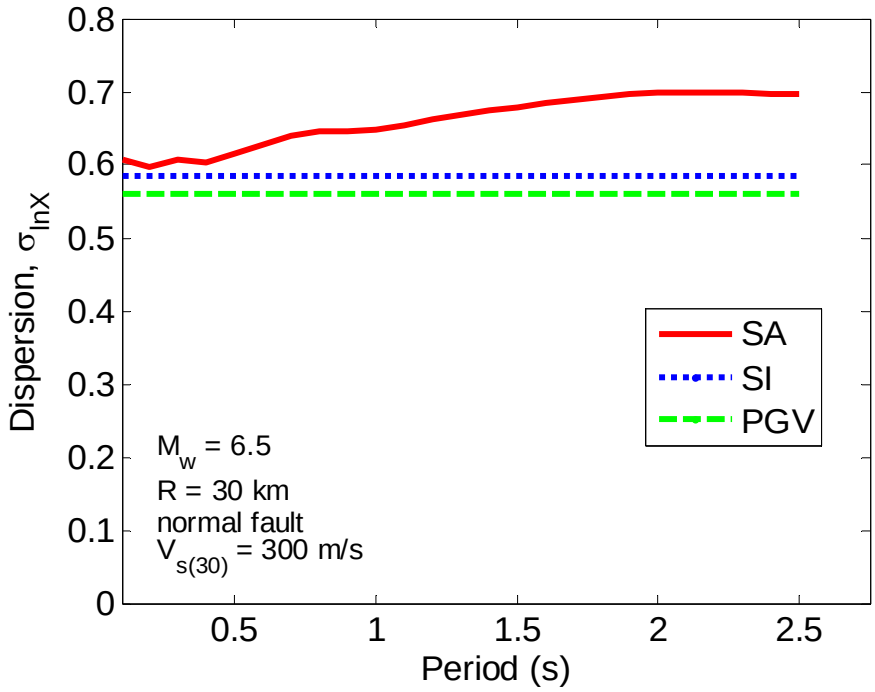


Figure 12-6: Comparison of the magnitude of the lognormal standard deviation (dispersion) for spectral accelerations at various periods, *Sa*, peak ground velocity, *PGV*, and spectrum intensity, *SI*.

12.6.3 Variation in dispersion with magnitude and distance

In addition to the dispersion in the spectral acceleration ground motion prediction equations being a function of the vibration period, historically dispersion in ground motion prediction equations were also a function of earthquake rupture magnitude [e.g. 16, 17], with the dispersion reducing as the magnitude increased; but were generally assumed to be independent of source-to-site distance. However, in the recent release of the NGA ground motion equations [e.g. 12, 13, 14], some prediction equation developers appear to have decided that in the revised models, the magnitude dependence is insignificant, and therefore have been determined independent of rupture magnitude.

Because the computation of *SI* proposed here does not enforce any dependence/independence of the dispersion on magnitude or distance, it is possible that the *SI* prediction is dependent on magnitude and distance, even if the *Sa* prediction is not. Figure 12-7 illustrates the variation in dispersion as a function of both magnitude and (Boore-Joyner) distance using the Boore and Atkinson [13] NGA prediction equation. Figure 12-7a illustrates that the dispersion is observed to increase as the (moment) magnitude increases, the opposite trend in comparison to the magnitude-dependent dispersion for spectral acceleration terms. This increasing trend is however relatively insignificant, providing an increase from 0.55 for $M_w = 5$ to 0.60 for $M_w = 8$ (8.3% increase) in comparison to the variation in dispersion of 0.38-0.545 (30.3% increase) from $M_w = 5$ to 7.4 used in Campbell and Bozorgnia [17]. Figure 12-7b illustrates the dependence of the dispersion on source-to-site distance for the particular fault rupture scenario considered. In this case, there is evidently no dependence on the source-to-site distance with the dispersion ranging from approximately 0.58-0.587.

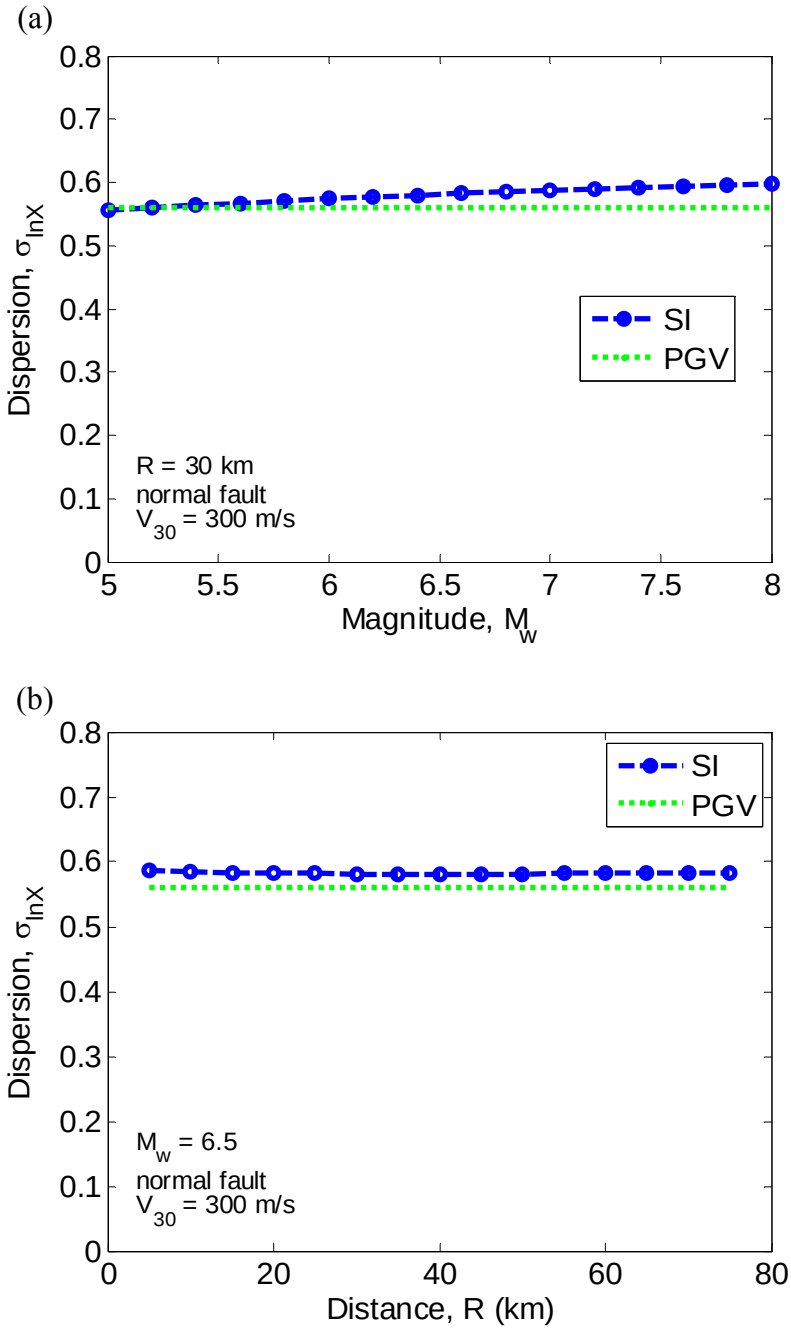


Figure 12-7: Variation in lognormal standard deviation (dispersion) as a function of: (a) magnitude; and (b) source-to-site distance.

12.7 Conclusions

Seismic response analysis of structures requires the use of ground motion intensity measures which both *efficiently* predict the seismic response of structure due to ground motion excitation, and are *predictable* in the sense that a ground motion prediction equation exists which can estimate the IM with a relatively low uncertainty for a given earthquake rupture scenario. In this chapter a ground motion prediction equation for spectrum intensity (*SI*),

defined as the integral of the pseudo-velocity spectra over the 0.1-2.5s period range, has been proposed based on spectral acceleration (*Sa*) prediction equations, which are both extensively developed and abundant in research publications and earthquake engineering practice. The theoretical basis behind the development of the *SI* prediction equation based on *Sa* prediction equations was presented and various pertinent issues relating to its development were addressed. In particular, it was shown that the lognormal distribution can be used to accurately capture the distribution of *SI* for a given earthquake scenario both over the 'body' and 'tails' of the empirical distribution. The computation of *SI* based on *Sa* prediction equations can be adequately performed using a discretization of the vibration period of less than 0.2 seconds. The proposed approach was verified via comparison with direct prediction equations for *SI*. It was also illustrated that because *Sa* terms at various periods are not perfectly correlated, *SI* is a highly *predictable* IM as compared to other common IMs such as *Sa* and *PGV*.

12.8 Acknowledgements

The first author would like to acknowledge funding from the New Zealand Tertiary Education Commission. Fruitful discussions of the first author with Dr Laurentiu Danciu and Dr Jack Baker during the initial development of this work, and Mr. Nirmal Jayaram near its completion are greatly appreciated.

12.9 References

[1] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[2] Kramer SL and Mitchell RA. Ground motion intensity measures for liquefaction hazard evaluation. Earthquake Spectra 2006; 22(2): 413-438.

[3] Bradley BA, Cubrinovski M, and Dhakal RP. Performance based seismic response of pile foundations. Geotechnical Earthquake Engineering and Soil Dynamics IV (GSP181), Sacramento, CA, 2008.

[4] Bradley BA, Cubrinovski M, Dhakal RP, and MacRae GA. Intensity measures for the seismic response of pile foundations. Soil Dynamics and Earthquake Engineering 2009; 29(6): 1046-1058, DOI: 10.1016/j.soildyn.2008.12.002.

[5] Housner GW. The behaviour of Inverted Pendulum Structure During Earthquake. Bulletin of the Seismological Society of America 1963; 53(2): 403-417.

[6] Housner GW. Spectrum intensities of strong-motion earthquakes, in Symposium on earthquakes and blast effects on structures, Los Angeles, CA, 1952.

[7] Jayaram N and Baker JW. Statistical tests of the joint distribution of spectral acceleration values. Bulletin of the Seismological Society of America 2008; 98(5): 2231-2243, DOI: 10.1785/0120070208.

[8] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[9] Baker JW and Cornell CA. Correlation of Response Spectral Values for Multi-component Ground motions. Bulletin of the Seismological Society of America 2006; 96(1): 215-227.

[10] Inoue T and Cornell CA. Seismic hazard analysis of multi-degree-of-freedom structures. Stanford, CA, 1990. 70.

[11] Baker JW and Jayaram N. Correlation of spectral acceleration values from NGA ground motion models. Earthquake Spectra 2008; 24(1): 299-317.

[12] Abrahamson NA and Silva WJ. Summary of the Abrahamson & Silva NGA ground motion relations. Earthquake Spectra 2008; 24(1): 67-97.

[13] Boore DM and Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s. Earthquake Spectra 2008; 24(1): 99-138.

[14] Campbell KW and Bozorgnia Y. Campbell-Bozorgnia NGA horizontal ground motion model for PGA, PGV, PGD and 5% damped linear elastic response spectra. Earthquake Spectra 2008; 24(1): 139-171.

[15] Danciu L and Tselentis GA. Engineering Ground motion Attenuation Relationships for Greece. Bulletin of the Seismological Society of America 2007; 97(1B): 162-183, DOI: 10.1785/0120040087.

[16] Abrahamson NA and Silva WJ. Empirical response spectral attenuation relations for shallow crustal earthquakes. Seismological Research Letters 1997; 68(1): 94–126.

[17] Campbell KW and Bozorgnia Y. Near-source attenuation of peak horizontal acceleration from worldwide accelerograms recorded from 1957 to 1993, in 5th U.S. National Conference on Earthquake Engineering, Berkley, California, 1994, 283-292.

13.Intensity Measures for the Seismic Response of Pile Foundations

Bradley BA, Cubrinovski M, Dhakal RP, MacRae GA. Intensity measures for the seismic response of pile foundations. *Soil Dynamics and Earthquake Engineering* **29**(6): 1046-1058.

13.1 Abstract

In this study the efficacy of various ground motion intensity measures for the seismic response of pile foundations embedded in liquefiable and non-liquefiable soils is investigated. A soil-pile-structure model consisting of a two layer soil deposit with a single pile and a single-degree of freedom superstructure is used in a parametric study to determine the salient features of the seismic response of the soil-pile-structure system. A suite of ground motion records scaled to various levels of intensity are used to investigate the full range of pile behaviour, from elastic response to failure. Various intensity measures are used to inspect their efficiency in predicting the seismic demand on the pile foundation for a given level of ground motion intensity. It is found that velocity-based intensity measures are the most efficient in predicting the pile response, which is measured in terms of maximum curvature or pile head displacement. In particular, velocity spectrum intensity (*VSI*), which represents the integral of the pseudo-velocity spectrum over a wide period range, is found to be the most efficient intensity measure in predicting the seismic demands on the pile foundation. *VSI* is also found to be a sufficient intensity measure with respect to earthquake magnitude, source-to-site distance, and epsilon, and has a good predictability, thus making it a prime candidate for use in seismic response analysis of pile foundations.

13.2 Introduction

In recent strong earthquakes, a large number of pile foundations of high-rise buildings, bridge piers and storage tanks have been severely damaged, thus affecting the functionality of modern engineering structures designed according to the most advanced seismic standards [1, 2]. The soil-pile-structure interaction during strong ground shaking is a complex phenomenon involving a significant change in stiffness and strength of soils and large lateral loads on piles. Both inertial loads from vibration of the superstructure and kinematic loads imposed by lateral ground displacements contribute toward pile damage. In the 1995 Kobe earthquake, for example, piles in non-liquefied deposits were subjected to peak cyclic ground displacements of about 30 cm whereas the respective displacements in liquefied soils reached about 40-50 cm causing damage to numerous pile foundations [3, 4]. When evaluating the seismic performance of pile foundations, one should estimate the inelastic response of piles with the ultimate goal being the assessment of the potential damage to the foundation and consequent effects on the safety and functionality of the superstructure.

Rigorous seismic assessment of pile foundations (and their superstructures) requires the use of a probabilistic framework to account for the numerous uncertainties in quantifying the performance. Contemporary performance-based earthquake engineering (PBEE) evaluation is typically defined based on the PEER performance assessment methodology [5, 6] and consists of four key steps, as outlined schematically in Figure 13-1: ground motion estimation, seismic response estimation, damage estimation and loss estimation. Each of these four steps is linked together through the use of intensity measures (IM*'*s), engineering demand parameters (EDP*'*s) and damage measures (DM*'s*) to give metrics of seismic performance, termed decision variables (DV's). When applied to a given structure, the methodology allows a probabilistic quantification of seismic performance using measures such as: annual rate of collapse, expected annual loss and annual rate of exceeding some economic loss, all of which can be easily communicated to decision makers and society.

Step 1 of the PEER methodology as given in Figure 13-1 is the use of a ground motion IM to provide a (probabilistic) relationship between the ground motion hazard at the site and the resulting seismic response. Clearly, the determination of seismic demands on piles (and structures in general) during strong ground motion is burdened by significant uncertainties. Such uncertainties may be either aleatory (randomness) or epistemic (knowledge-based) in nature. Obviously, one endeavours to minimise such uncertainties in order to improve system

reliability and reduce seismic risk. Record-to-record (RTR) randomness (i.e. variation in seismic demand due to different ground motions of the same intensity) is a key uncertainty in the relationship between seismic demand and seismic intensity [7]. RTR randomness results from the complex mechanism of source rupture, wave propagation and scattering, local site and soil-structure interaction effects, as well as quantifying the ground motion shaking using a scalar or vector IM. It is therefore necessary to identify IMs which display a good correlation with the performance metric to be evaluated, thus reducing this uncertainty.

In this chapter intensity measures for the seismic response of pile foundations embedded in both liquefiable and non-liquefiable soils is investigated. Nonlinear dynamic analyses of a soil-pile-structure model are used to identify the most appropriate EDP and IM for pile foundations. Firstly, consideration is given to the determination of an effective EDP that describes the seismic response of the pile and its associated damage. Next, various candidate IMs are examined and ranked based on their resulting aleatory uncertainty in predicting EDP (i.e. efficiency [8]); their response prediction independence on conventional measures of ground motion (i.e. sufficiency [9]); and the uncertainty in predicting the IM itself (i.e. predictability [10]). Demand hazard curves are used to illustrate the benefits of appropriately selecting IMs. Finally, effects of soil-pile-structure interaction and liquefaction on the deformation mechanism of the system are also discussed.

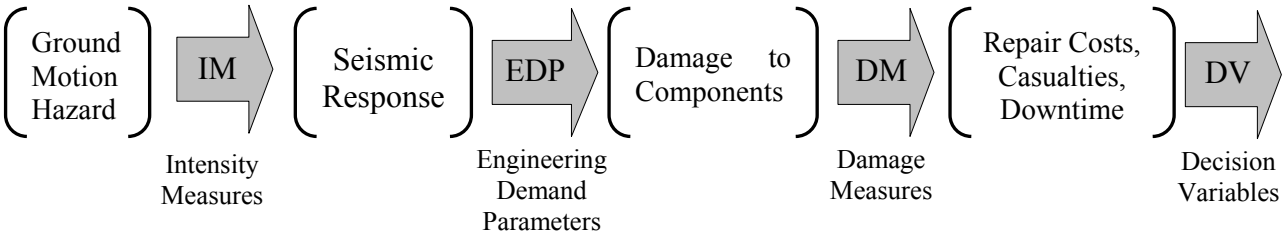


Figure 13-1: Schematic illustration of the key tasks in the PBEE methodology

13.3 Adopted soil-pile-structure model

A conceptually simplified numerical model is adopted in this investigation to adequately capture the salient features of the pile response without onerous complexity. As shown schematically in Figure 13-2a the model consists of a two-layer non-liquefiable soil deposit with a single pile, rigid pile cap and a superstructure represented by a single degree-of-freedom. To investigate the effects of different material and geometrical properties, six different scenarios are considered as summarised in Figure 13-2b (herein for brevity 'S' will be used as shorthand notation for 'scenario'). S1-S4 involve variations of soil stiffness and

strength, pile properties, weight and period of the superstructure, all of which are based on typically observed configurations in engineering practice. S5 and S6 are the same as S1 and S3, respectively, but with no superstructure in the FE model. S5 and S6 therefore represent cases where the response of the pile is entirely due to kinematic effects of the imposed ground deformations, without inertial effects from the superstructure vibration.

For each of the scenarios, the weight of the SDOF superstructure (in the FE model) was based on a typical axial load applied to a single pile within a pile group of a representative superstructure. The stiffness of the SDOF superstructure was then computed to achieve the same target period as that of the representative superstructure. A rigid pile cap with a width of 3m was modelled using an elastic beam element with a large stiffness that effectively provided a 'fixed' condition at the head of the pile. The FE code DIANA-J [11] was used in the analyses, which allows for fully-coupled effective-stress analysis of the soil medium capturing pore water pressure development, its redistribution and dissipation. An elastic-plastic constitutive model for sand, the Stress-Density (S-D) model [12, 13] was used for modelling the soil behaviour. Two series of analyses were performed. In the first series, seismic analyses of the scenarios given in Figure 13-2b were performed with the S-D model parameters set such that pore-pressure build-up is effectively suppressed to simulate non-liquefiable soil. In the second series of analyses, the effects of excess pore water pressures and eventual liquefaction were considered. The S-D model has been extensively verified through rigorous simulations of down-hole array records at liquefaction sites, seismic centrifuge tests, large-scale shake-table tests on pile foundations and case histories of damaged piles in the 1995 Kobe earthquake (e.g. [14] and references therein).

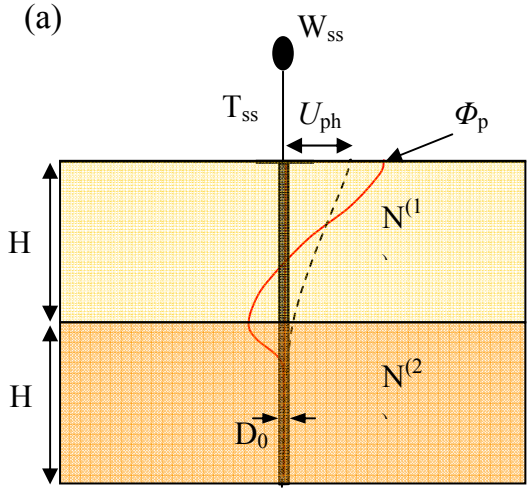
Conventional shear modulus reduction curves for sand proposed by Seed and Idriss [15] were used to determine model parameters controlling the nonlinear stress-strain behaviour of the soil (Figure 13-2c). Figure 13-2d illustrates the adopted liquefaction resistance curves for the two different soil layers considered. The cyclic stress ratio (CSR) at 20 cycles for the $N^{(1)}$=10 and $N^{(1)}$=20 soils were 0.15 and 0.26, respectively. Dilatancy parameters of the constitutive model used in the analyses were based on those of Toyoura sand [13] and then modified to achieve the liquefaction resistance depicted in Figure 13-2d.

The pile was modelled using one-dimensional beam elements with a hyperbolic moment-curvature relationship as an approximation to the M-ϕ relationship obtained from fibre-based section analyses (Figure 13-2e; Table 13-1). Similarly, nonlinear beam-elements with stiffness and mass defined by the parameters shown in Figure 13-2b were used to model the superstructure.

Table 13-1: Properties of pile foundations used in analysis.

	D_0 = 0.4m PC pile	D_0 = 1.2m RC pile
Initial Stiffness, EI (MN-m^2)	42.3	2340
Ultimate Moment, M_u (MN-m)	0.135	2.46
Axial Load, N, (kN)	400	2500
Reinforcement ratio, ρ	0.3%	0.5%
Prestress level (MPa)	4.0	0.0
Steel yield stress, f_y (MPa)	1320	340
Concrete crushing strength, f_c' (MPa)	50.0	25.0

(a)

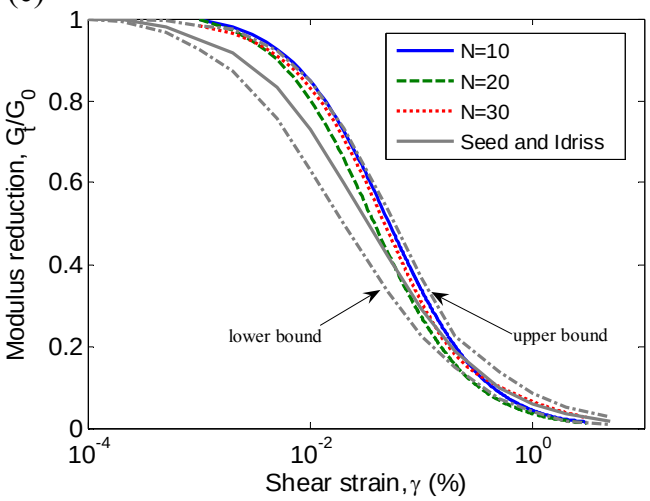


(b)

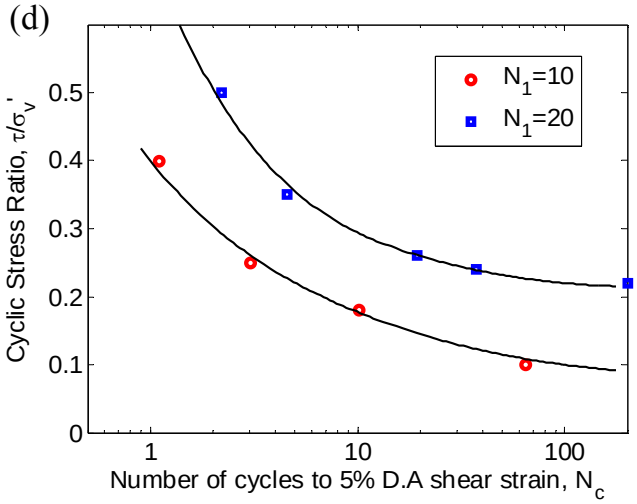
Model	Scenario					
Parameter	**1**	**2**	**3**	**4**	**5**	**6**
$N^{(1)}$	10	20	10	20	10	10
$N^{(2)}$	30	30	30	30	30	30
W_{ss} (kN)	2500	2500	400	400	-	-
T_{ss} (s)	1.8	1.8	0.8	0.8	-	-
D_0 (cm)	120	120	40	40	120	40
H (m)	10	10	10	10	10	10

$N^{(i)}$ = SPT blow count of layer *i*; W_{ss} = weight of superstructure
T_{ss} = period of superstructure; D_0 = diameter of pile
H = depth of soil layer

(c)



(d)



(e)

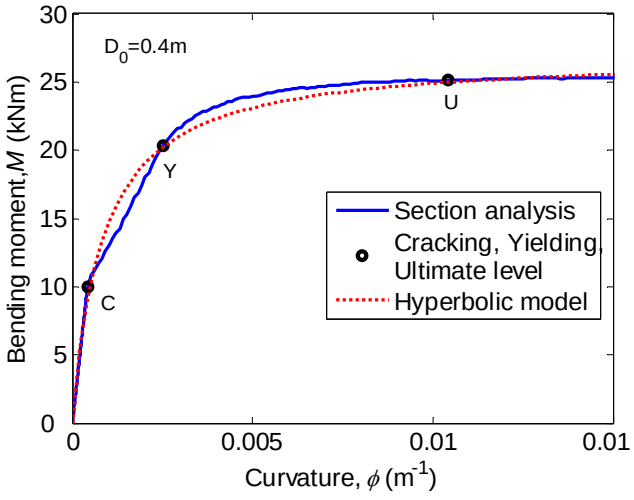


Figure 13-2: Soil-pile-structure model used: (a) schematic illustration of model; (b) scenarios considered; (c) modulus reduction curves; (d) liquefaction resistance curves; and (e) hyperbolic approximation of M-Φ relationship for the pile.

13.4 Ground motions and intensity measures investigated

A suite of 40 ground motion records compiled by Medina and Krawinkler [16] were used as input motions in the non-linear time history analyses. This suite contains ground motions recorded on stiff soils with magnitude and distance ranges of M_w = 6.5-6.9 and R = 13.3-39.3 km, respectively. The suite is termed 'ordinary' by Medina and Krawinkler, as none of the records show effects of near-fault motions (i.e. directivity or 'fling' effects), and all ground motions were recorded on stiff soils. In the series of analyses with non-liquefiable soils the FE model was subjected to a base input motion scaled to ten different levels of peak ground accelerations between 0.1 and 1.0g in steps of 0.1g. In the series of analyses with liquefiable soils the input motions were scaled from 0.05g to 0.5g in steps of 0.05g. The smaller intensity increments for liquefiable soils were used to follow the extent of liquefaction with intensity more accurately. Thus, using the forty different ground motion records as input motions, a total of 400 analyses were performed for each of the six scenarios listed in Figure 13-2b (for both liquefiable and non-liquefiable analyses).

Significant research over the past decade has focused on determining optimal IMs for predicting structural response due to earthquake excitation (e.g. [8, 9, 17, 18]). Clearly, different ground motion parameters should be used as IMs depending on the governing deformation mechanisms and response characteristics of the structure considered. Since pile foundations involve both kinematic and inertial effects due to soil and superstructure response, respectively, it is necessary to examine a wide range of potential IMs and identify the optimum IM for prediction of the pile response. The determination of an optimal IM for prediction of a level of seismic demand is guided by the concepts of 'efficiency' [8], 'sufficiency' [9], and 'predictability' [10]. Efficiency is related to the aleatory uncertainty in the *EDP*|*IM* relationship; sufficiency to the independence of the response to ground motion selection; and predictability to the uncertainty in predicting the ground motion IM using a ground motion prediction equation.

A total of 19 different IMs are considered for correlation with the pile response and are summarised in Table 13-2. All IMs were computed from the base input ground motion and IM definitions can be found in Riddell [19]. The investigated IMs include acceleration-, velocity-, and displacement-based ground motion intensity measures, both of peak and cumulative nature.

Table 13-2: Intensity Measures used in analyses.

ID	Intensity Measure (IM)
1	Peak ground acceleration, *PGA*
2	Peak ground velocity, *PGV*
3	Peak ground displacement, *PGD*
4	Significant duration, D
5	*PGV*/*PGA*, V_{max}/A_{max}
6	RMS acceleration, RMS_a
7	RMS velocity, RMS_v
8	RMS displacement, RMS_d
9	Arias intensity, I_a
10	Specific energy density, dE
11	Cumulative absolute velocity, CAV
12	Acceleration spectrum intensity, ASI
13	Velocity spectrum intensity, *VSI*
14	Sustained maximum acceleration, SMA
15	Sustained maximum velocity, SMV
16	Spectral acceleration, S_a(T=0.4s,5%)
17	Spectral acceleration, S_a (T=0.6s,5%)
18	Spectral acceleration, S_a (T=0.8s,5%)
19	Spectral acceleration, S_a (T=1.8 s,5%)

It is well known that the destructive potential of a ground motion is dependent on its intensity, frequency content and duration. Thus, each of the ground motion IMs presented in Table 13-2 quantifies some or all of these characteristics of the ground motion. For example, peak quantities such as *PGA*, *PGV* and *PGD* account for the ground motion intensity only. Spectral acceleration terms account for ground motion intensity at a specific period of vibration, while *ASI* and *VSI* account for the spectral quantities over a relevant range of vibration periods. *CAV* accounts for both intensity and duration, while Arias intensity, $I_{a,}$ accounts for all three ground motion characteristics. It is worth noting that the spectral acceleration IMs at periods of 0.8 and 1.8s are selected to coincide with the first-mode period of the superstructure considered.

13.5 Measure of seismic demand on piles

The first question to be answered before determining which ground motion IMs correlate well with seismic demand on piles is: how is the seismic demand measured? Ideally, the adopted engineering demand parameter (EDP) would correlate perfectly with the occurrence of damage in the pile. In comparison to the research attention that the *EDP*|*IM* relationship has received, little research has focused on determining optimal EDPs which

correlate well with damage measures (DMs) in structural components and pile foundations in particular.

As with any engineering material, the seismic demand on a pile foundation is generally related to the hysteretic energy that is dissipated during ground motion excitation. Hysteretic energy is typically expressed as a function of both peak and cumulative deformations, one typical example being the damage index of Park and Ang [20]. However, due to the difficulty of developing fragility relationships between complex EDPs and damage in components, simple measures of demand are typically adopted. In the performance-based assessment of structural systems, typical EDPs used are cumulative plastic rotation and peak interstorey drift [8, 21]. For simplicity, only peak response quantities will be used to quantify the seismic demand on the pile foundation in this study.

The peak pile curvature, ϕ_p, would seem the most obvious candidate to use for pile demand, as it directly relates to the peak strains at the critical section of the pile and hence the extent of damage. However, peak pile curvature, being a localised measure of demand, is not easily related to the global response of the pile and the foundation system. The peak lateral displacement of the pile head (i.e. at the pile cap), U_{ph}, is such an EDP which represents the global response of the foundation. In addition, the peak foundation displacement can be used as a proxy for damage to connections, and post-earthquake serviceability of relevant lifelines (e.g. electricity and water piping).

It is intuitive that the peak pile head displacement and the peak pile curvature are correlated based on the typical deformation pattern of piles caused by earthquake-induced lateral loads. Figure 13-3 illustrates the correlation between peak pile displacement, U_{ph}, and peak pile curvature, ϕ_p, obtained from the 400 non-linear finite element analyses on non-liquefiable soils for S1, S3, S5, and S6 as described in previous section. Here, each symbol represents the result of one analysis while the lines are best fit approximations of the trend. Note that the results of S2 and S4 are not shown here as they are similar to S1 and S3, respectively. Curvature results presented here have been obtained from the difference in the nodal rotations in the beam elements and were not found to be sensitive to the element length used (which was constant in all analyses). Figure 13-3a and 13-3b illustrate that the correlation for S3 is not as strong as that for S1. Note that the difference between S1 and S3 is in the dynamic characteristics of the superstructure and pile properties. S5 and S6 are the same as S1 and S3, respectively, but with no superstructure, and both show good correlation between the pile curvature and pile head displacement. The good correlation for S5 and S6

(the only difference between the scenarios being the properties of the pile) indicates that the relatively poor correlation for S3 cannot be explained by the pile properties or its flexibility. Therefore the reduction in the correlation between curvature and pile head displacement for S3 (and S4 for that matter) is apparently due to significant inertial effects acting from the vibration of the superstructure on the response of the pile head. These inertial effects cause relatively large curvatures in the pile in a localised region just below the pile cap and are more pronounced for relatively flexible piles.

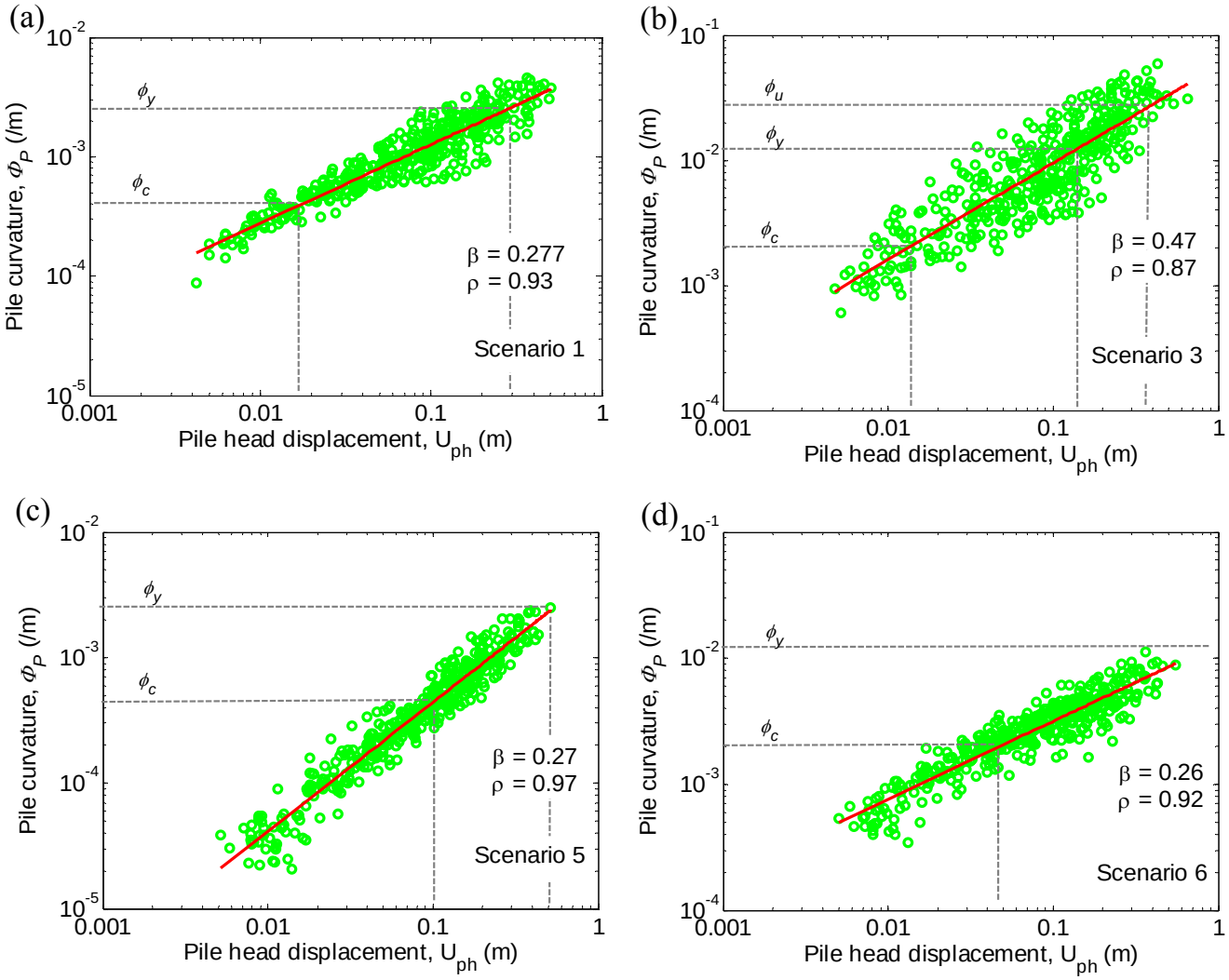


Figure 13-3: Correlation between peak pile displacement and peak pile curvature from 400 nonlinear FE analyses for: (a) scenario 1; (b) scenario 3; (c) scenario 5; and (d) scenario 6.

Figure 13-4 illustrates the mean relationship between pile head displacement and pile curvature for all six scenarios (i.e. the regression lines in Figure 13-3). Firstly, it is evident that the relationships for the relatively flexible 0.4m diameter pile shown in Figure 13-4a all lie above the respective relationships for the relatively stiff 1.2m diameter pile shown in

Figure 13-4b. In other words, for a given level of pile head displacement, U_{ph}, the peak curvature, ϕ_p, of the flexible pile is, on average, larger than that for the stiffer pile. The ratio of curvatures for a given displacement is approximately eight for large curvatures, and approximately five for small (essentially elastic) curvatures Assuming that: (i) the plastic hinge length is proportional to the pile diameter; (ii) the effective length of the pile is constant; and (iii) that the pile cap provides zero rotation at the pile head, it can be deduced that for a given displacement at the pile head, the inelastic curvature for the 0.4m diameter pile will be three times that of the 1.2m diameter pile. However for the soil-pile model considered in this study, for a given pile-head displacement the curvature in the stiff pile will be further reduced (compared to the flexible pile) by the fact that the effective pile length will be longer for the stiff pile, and also because the larger relative stiffness between the pile and pile cap (compared to the flexible pile) will result in a larger rotation of the pile cap. Figure 13-4 also shows that the soil stiffness (within the range of values examined in this study) appears to have little effect on the relationship between pile head displacement and pile curvature, as evident from the results for scenarios S1, S2 (1.2m diameter pile with $N^{(1)} = 10$ and 20, respectively) and S3, S4 (0.4m diameter pile with $N^{(1)} = 10$ and 20, respectively). The relative insensitivity of the pile response on the soil SPT blowcount can be explained by the fact that when changing the SPT blowcount from 10 to 20 the relative stiffness only changes by approximately 20% [22]. In contrast, when using the properties of the piles given in Table 13-1, it can be shown that the difference in the relative stiffness between the pile and soil for the two different piles is a factor of about 2.3. Finally, the large deviation of the results for scenarios 5 and 6 illustrates the significant effect of the inertial load from the vibration of the superstructure on the pile response.

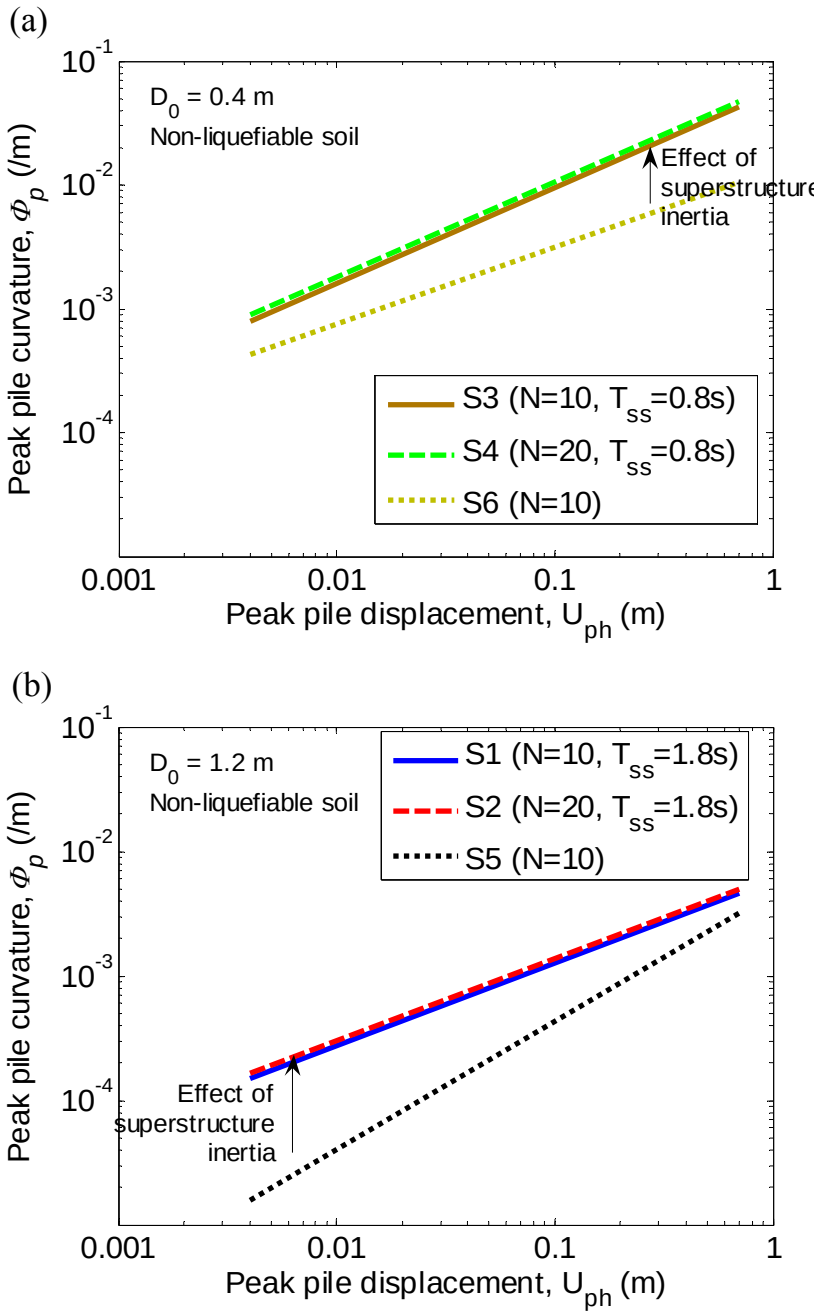


Figure 13-4: Illustration of the variation in the relationship between peak pile curvature and peak pile head displacement for different pile and soil conditions: (a) 0.4m diameter pile; and (b) 1.2m diameter pile.

Clearly, a unique relationship between pile curvature (which relates directly to pile damage) and pile head displacement does not exist, but rather this relationship is a function of the relative stiffness of the pile and the soil, and the presence of inertial loads at the pile head. As illustrated in Figure 13-3, for a single scenario however, there exists a relatively strong correlation between the pile head displacement, U_{ph}, and pile curvature, ϕ_p. As a result of this high correlation between peak pile curvature and peak pile head displacement for a single

scenario, those IMs which efficiently predict the peak pile head displacement (as discussed in the next section) should also be efficient in predicting the peak pile curvature. This was found to be valid for all of the intensity measures investigated in the following sections. Thus for brevity, discussion herein on the pile response is given in terms of (peak) pile head displacement only, but readers should interpret the results as also applying to peak pile curvature.

13.6 Intensity measures for non-liquefiable soils

13.6.1 Efficiency

Figure 13-5a and 13-5b illustrate the observed pile head displacements from the 400 analyses in S1 for two intensity measures, peak ground displacement (*PGD*) and velocity spectrum intensity (*VSI*), respectively. The plots indicate the efficiency of these IMs. It is apparent that there is a significantly reduced scatter in the relationship between U_{ph} and *VSI* ($\beta = 0.25$) as compared to that of U_{ph} and *PGD* ($\beta = 0.54$). If a 90% confidence interval (CI) [23] is constructed for the median level of demand for a given level of intensity, the ratio between the 5% and 95% values is 5.9 for IM = *PGD*, and 2.3 for IM = *VSI* (i.e. the 95% value is 2.3 times the 5% value). Hence, the reduced dispersion gives a significantly reduced CI for the median demand. Obviously, in addition to this reduced CI for the median, a small dispersion means a smaller probability of very large demands for a given intensity, which affects the demand hazard [10].

Figure 13-6 summarises the dispersion values in the prediction of the pile head displacements for the 19 examined IMs for scenarios S1-S4 in non-liquefiable soils. There are several key points to note in Figure 13-6. With the exception of *PGV* for S1 and S2, the peak response values (i.e. *PGA*, *PGV*, *PGD*) tend to poorly correlate with the pile head displacement. Similar trends are observed with the root-mean-square (RMS) IMs. In Figure 13-6, the difference (in dispersion values) between S1 and S2, and S3 and S4 for the same IM is relatively insignificant. This suggests that changing the SPT blow count of the soil from $N = 10$ to $N = 20$ (i.e. for S1 and S3, $N = 10$; while S2 and S4 have $N = 20$) does not affect the efficiency of various IM*s* (recall that the soil stiffness made little difference in the relationship between pile head displacement and pile curvature also). This is not the case for liquefiable

soils as discussed later in the chapter.

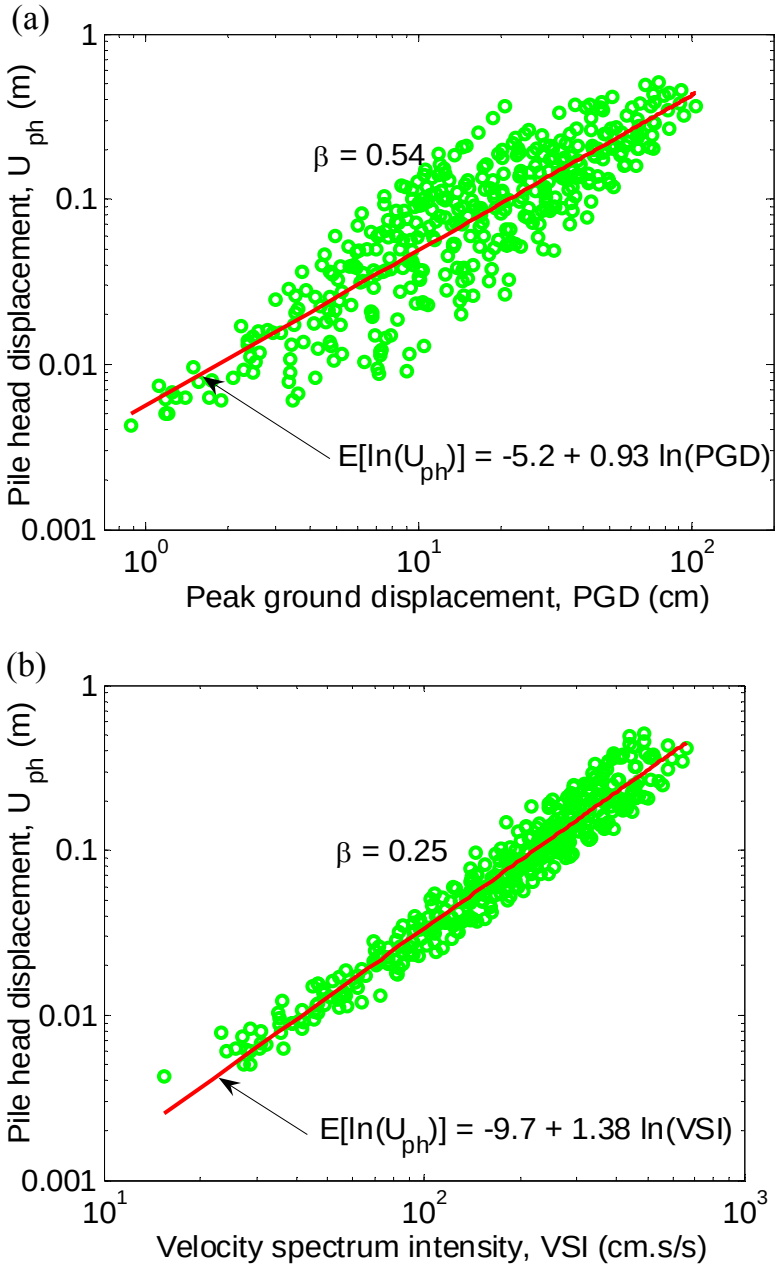


Figure 13-5: Comparison of EDP-IM scatter plots for: (a) Peak ground displacement, *PGD*; and (b) Velocity spectrum intensity, *VSI*, for the 400 analyses in scenario 1.

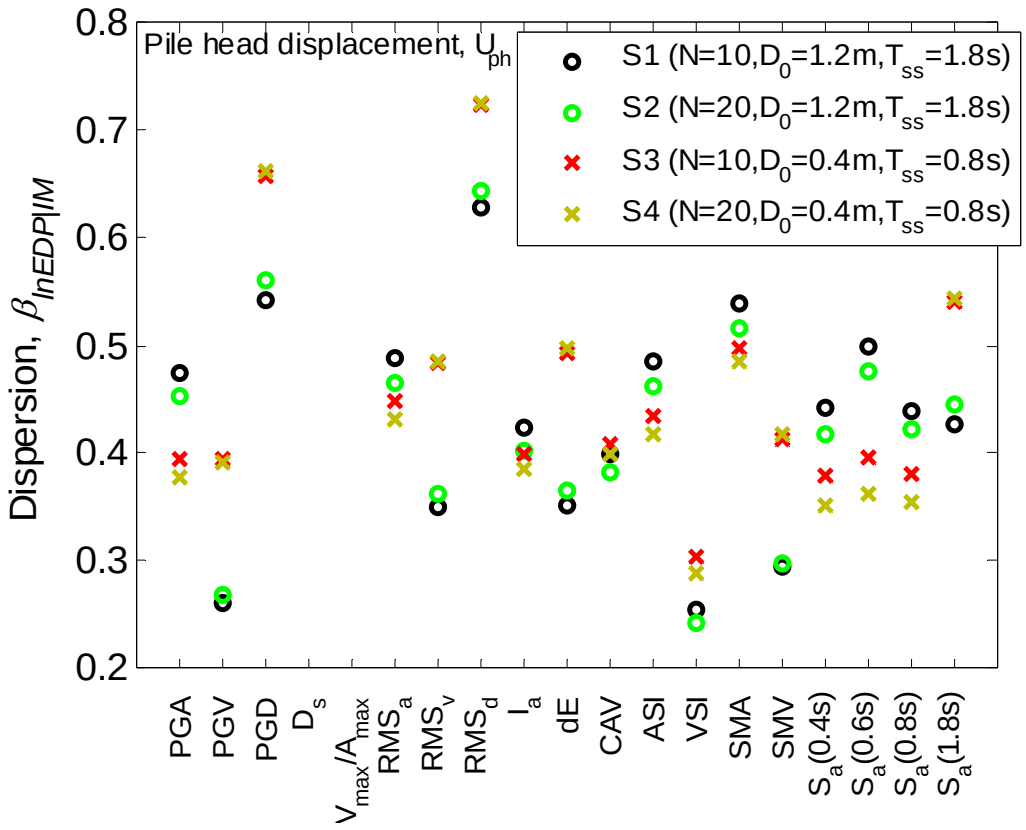


Figure 13-6: Efficiency of the candidate IMs for the four base scenarios (non-liquefiable soils).

The effect of the superstructure properties on the efficiency of the various IMs is however quite significant. Recall that S1 and S2 have a 1.8s period superstructure, while S3 and S4 have a 0.8s superstructure. This effect on IM efficiency is particularly pronounced for peak IMs such as *PGA* and *PGV*, where it can be seen using *PGA* is efficient for S3 and S4 (relative to S1 and S2), while *PGV* is significantly more efficient for S1 and S2 (relative to S3 and S4). In fact, all the velocity-based IMs (*PGV*, RMS_v, *CAV*, *VSI*, *SMV*) are observed to efficiently predict the pile head displacement when the superstructure period is 1.8 seconds, as compared to the 0.8 second superstructure. The converse is true for all the acceleration-based IMs (*PGA*, RMS_a, *ASI*, *SMA*, $S_a(T = 0.4,\ 0.6,\ 0.8\text{s})$), except of course $S_a(T = 1.8\text{s})$. These observations are in line with those of Riddell [19], who noted that the seismic response of stiff systems (with a low period of vibration) are efficiently predicted by acceleration-based IMs, while more flexible systems (in the constant velocity region of the response spectra) are better predicted by velocity-based IMs.

All four spectral acceleration IMs used were moderately efficient in predicting the pile head displacements. For S3 and S4 (when the superstructure period was 0.8s), and for S1 and

S2 (when the superstructure period was 1.8s) $S_a(T$=0.8s) and $S_a(T$=1.8s), respectively result in dispersions in the range of 0.36-0.44. The fact that the pile response is dependent on both kinematic and inertial mechanisms from the ground displacements and superstructure acceleration, respectively, is a likely reason that using the IMs at the vibration period of the superstructure (thus specifically targeting the inertial effects) does not result in a highly efficient IM. $S_a(T$=0.4s), $S_a(T$=0.6s) and $S_a(T$=0.8s) were used as a proxy for the kinematically-induced response of the pile. The moderate performance of these spectral terms (which have dispersions of 0.35-0.5) can be attributed to the fact that unlike structural systems (which exhibit a stable initial vibration period until significant inelastic action occurs due to yielding), no such initial period of soil exists during strong ground motion with inelastic behaviour occurring at low strain levels (e.g. Figure 13-2c).

In all four scenarios shown in Figure 13-6, *VSI* has the smallest dispersion. In other words , *VSI* has the best correlation with the pile response and provides the best prediction of the pile head displacement. *VSI* (also referred to as Housner's intensity, *SI*, but referred to as *VSI* here to clearly distinguish between *ASI*) is defined as [24]:

$$VSI = \int_{0.1}^{2.5} PSV(T,5\%)\mathrm{d}T \tag{13-1}$$

where $PSV(T,5\%)$ = the pseudo-velocity spectral ordinate for a period of vibration of T and 5% viscous damping. It can be seen from Equation (13-1) that unlike *Sa*, *VSI* uses information about the properties of the ground motion record at multiple vibration frequencies, which appears to be significant in quantifying the complex and highly non-linear soil-pile-structure response.

13.6.2 Sufficiency

IM sufficiency is important in addition to efficiency in seismic response analyses. As seismic response estimation is based on the response to a finite number of ground motion records, then the distribution of earthquake parameters (e.g. source distance, magnitude) of the selected ensemble of ground motion records may not strictly represent the distribution of ground motions which may occur at the site in the future. It is therefore desirable that the results of the response analyses are practically independent (over a reasonable range) of these earthquake characteristics. If independence is not satisfied then multiple integration would be required over all significant earthquake parameters, significantly complicating the use of the

IM. A typical example of an IM not being independent of other earthquake parameters, is the dependence of liquefaction development on both *PGA* and moment magnitude, M_w [10].

In order to investigate the additional information that ground motion parameters provide for responses which are already conditioned on the primary IM (e.g. *VSI*), the dependence of the residuals of the responses on other ground motion parameters is investigated. The residual is the arithmetic difference between an individual data point, (EDP_i,IM_j), and the predicted EDP value from the regression, $(\overline{EDP}, IM_j)$. By taking the residual of the responses conditioned on the primary IM, the influence of this IM is removed. Thus, if a significant dependence is found between the residuals and some other ground motion parameter, it indicates that this ground motion parameter should also be considered when conducting the seismic response analyses. Using an F-test [23], a so-called p-value is used as a quantitative measure of this independence, which gives, in this case, the probability that the (absolute) slope of the trend line through the residuals is at least as large as that observed (and therefore that the EDP prediction based on IM is independent of the considered ground motion parameter). Although somewhat subjective, generally a p-value of less than 0.05 indicates that there is evidence that the slope is non-zero, and a p-value less than 0.01 indicates strong evidence.

Figures 13-7a-13-7f illustrate the sufficiency of *PGA*-, *PGV*-, and *VSI*-based pile displacement responses with respect to earthquake moment magnitude (M_w) and source distance (R) for the S1 analyses. Note that when conducting conventional PSHA, integration is performed over three parameters, M_w, R, and ε (epsilon), hence the reason why the sufficiency with respect to these parameters is important. It can be seen that the *PGV*-based response is less dependent on both magnitude and source distance than the *PGA*-based response. Nevertheless, the p-values in Figure 13-7a-13-7d are notably low (only the dependence on distance of *PGV*-based responses is above 0.05). It appears somewhat surprising that the sufficiency of *PGV*-based responses is poor considering that the efficiency of *PGV* for the S1 analyses was very good. It should be noted however that the dependence on magnitude should be interpreted with some caution due to the relatively small variation in the magnitude values of the 40 ground motion records used here. Figure 13-7e and 13-7f illustrate the sufficiency of the *VSI* based pile responses with respect to M_w and R, respectively. The sufficiency exhibited is significantly superior to that of *PGV* and *PGA* for both M_w and R, and the p-values of 0.54 and 0.56, respectively, illustrate that practically there is no dependence on M_w or R when scaling records for pile response based on *VSI*.

Epsilon, ε, is defined as the number of standard deviations IM = im is from the median value, $\overline{IM}$, predicted from a ground motion prediction relationship. Epsilon is important in structural analyses as it has been shown that it is a proxy of the spectral shape in the frequency region of the IM, which is usually set equal to the fundamental period of the structure, $S_a(T_1)$ [17, 18]. The spectral shape will have a direct effect on the level of excitation which will occur due to high mode effects and effective period lengthening due to nonlinear behaviour. As *VSI* is dependent on the spectral velocity over a range of periods, then it directly accounts for the shape of the spectra (over this range of periods), and thus it seems logical that the dependence of *VSI* on epsilon is somewhat less compared to its effect when using $S_a(T_1)$. Figure 13-8 illustrates the sufficiency of *VSI* with respect to epsilon for S1. Although there appears to be a slight negative trend in the residuals as a function of epsilon, the p-value of 0.03 indicates that this dependence is relatively insignificant. Similar results for the sufficiency of *VSI* with respect to epsilon were observed for the other scenarios investigated in this study.

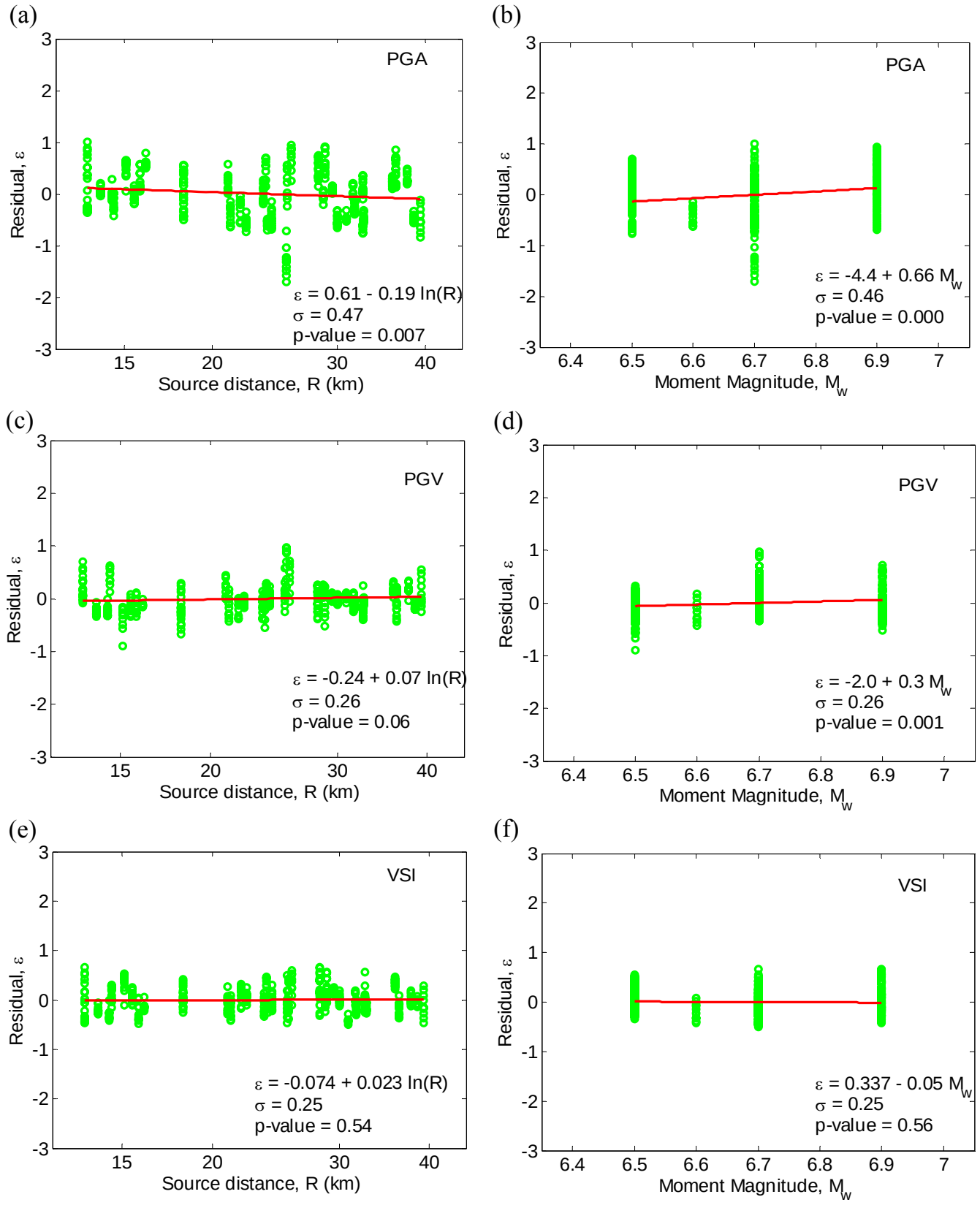


Figure 13-7: Sufficiency of peak ground velocity, peak ground acceleration and velocity spectrum intensity with respect to magnitude and distance.

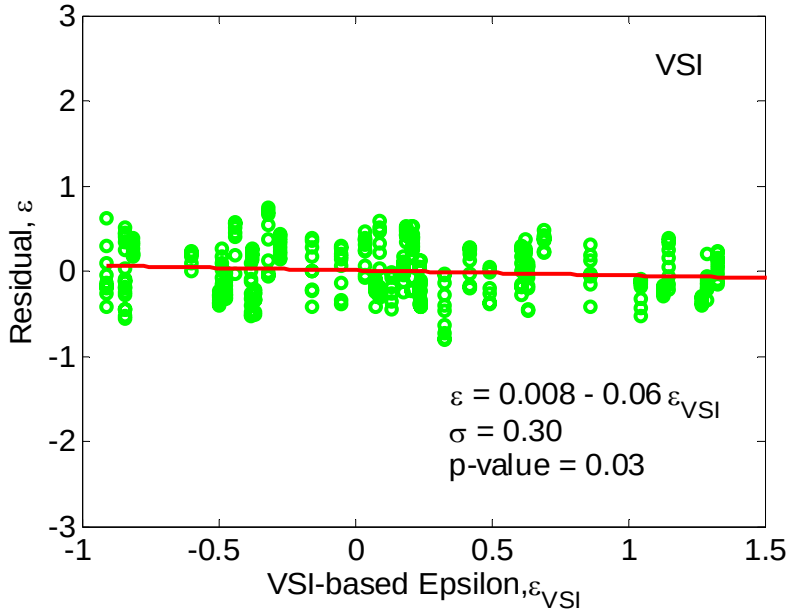


Figure 13-8: Sufficiency of *VSI* with respect to the ground motion parameter epsilon.

13.6.3 Distribution of demand for a given intensity

When considering the relationship between seismic demand and ground motion intensity (*EDP*|*IM*) within the PBEE framework, knowledge of the distribution of *EDP*|*IM* is required. While several investigations in structural earthquake engineering indicate that the lognormal distribution can accurately represent the distribution of *EDP*|*IM* (e.g. [25, 26]), the highly nonlinear soil-pile-structure interaction and potential occurrence of liquefaction in strong ground motion shaking may mean that the seismic demand on the piles and soil response in general does not follow this common analytical distribution.

Figure 13-9 illustrates the validity of the lognormal assumption for a variety of seismic intensity measures, scenarios and intensity levels. Namely, EDPs considered were peak free-field soil displacement (U_g), peak pile head displacement (U_{ph}), and peak pile curvature (ϕ_p) based on IMs: *PGA*, *PGV* and *VSI*. Figure 13-9 illustrates that various combinations of these EDPs and IMs all can be assumed to have a lognormal distribution for *EDP*|*IM* based on the Kolmogorov-Smirnov (K-S) goodness-of-fit test at a 10% confidence level [23]. From the above observations (as well as additional analyses not discussed herein) it would seem logical that the seismic demand will also be well approximated by the lognormal distribution for piles in liquefiable and non-liquefiable soils.

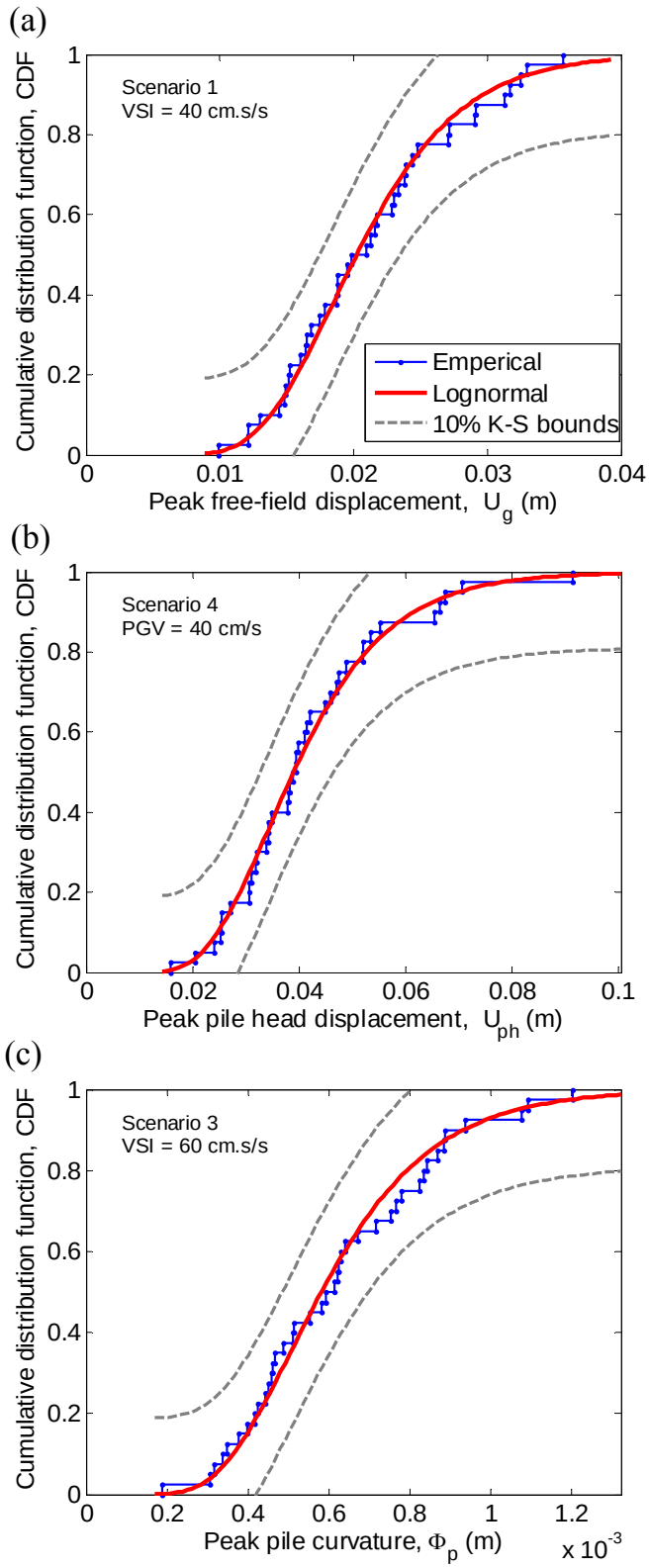


Figure 13-9: Goodness-of-fit tests for the distribution of: (a) peak free-field displacements; (b) peak pile head displacements; and (c) peak pile curvatures for various scenarios and levels of ground motion intensity measures.

13.6.4 Predictability

When selecting an IM suitable for seismic response analysis, consideration must also be given to the predictability of the IM which influences the ground motion hazard curve determined via PSHA. Kramer and Mitchell [10] illustrate how significant uncertainty in the ground motion prediction model for the IM will cause an increase in the ground motion hazard, and in turn an increase in the demand hazard. In other words, if the predictability of an IM is poor, then the accuracy in predicting the pile response (for a given earthquake scenario, M_w, R) will also be poor. In this context, it is important to examine the predictability of *VSI*, which from the results of the analyses presented so far, is undoubtedly the optimal IM for pile response in terms of efficiency and sufficiency.

Unlike *PGA*, for example, there are few ground motion prediction equations for directly predicting *VSI*. Here, the *VSI* prediction equation presented in Danciu and Tselentis [27] will be used. Although beyond the scope of this study, it is noted that it is possible to obtain a ground motion prediction equation for *VSI* from spectral acceleration prediction relationships [28], thus allowing *VSI* to be computed at all sites where applicable spectral acceleration relationships are available.

Comparison of the various ground motion prediction equations of Danciu and Tselentis [27] illustrates that *VSI* is also an IM with good predictability; the prediction equation for *VSI* has a dispersion of 0.73 compared to that of *PGA* and *PGV* which have dispersions of 0.67 and 0.71, respectively (in terms of the natural logarithm, lnIM, and not the base 10 logarithm, $\log_{10}$IM).

By combining the ground motion hazard with the seismic response analyses, the demand hazard can be obtained from:

$$\lambda_{EDP}(edp) = \int G_{EDP|IM}\left(edp \mid im\right)\left|\frac{d\lambda_{IM}(im)}{dIM}\right|\mathrm{dim} \tag{13-2}$$

where $G_{EDP|IM}\left(edp \mid im\right)$ is the complementary cumulative distribution function for *EDP*|*IM*, which gives the probability of exceeding EDP = *edp* given IM = *im*; and $\lambda_Z(z)$ gives the annual rate of exceeding $Z = z$. In this case, Equation (13-2) gives the annual frequency of exceeding various levels of peak pile head displacement. Figure 13-10 illustrates the demand hazard curves obtained for S3 using *VSI* and *PGV* as IMs. For consistency regarding site classification factors and faulting style, the ground motion attenuation models of Danciu and Tselentis [27] were used for both *VSI* and *PGV*. In producing the demand hazard curves

shown in Figure 13-10, a hypothetical site on stiff soil is considered (as used in [10]) which is R=30 km from a single strike-slip fault point source, with a magnitude recurrence relationship given by $\log_{10}\lambda_{M_w} = 4.0 - 0.8M_w$. The dispersions in the seismic response analyses for *VSI* and *PGV* were 0.305 and 0.395, respectively. It is evident that the difference in dispersion is the key contributor to the difference in the hazard curves, since the attenuation relationships of both *VSI* and *PGV* have similar dispersions of 0.73 and 0.71, respectively [27].

Figure 13-10a illustrates that using the *VSI*- and *PGV*-based scaling of the ground motion records gives pile head displacements of approximately 0.18 m and 0.215 m, respectively for the 10% in 50 year exceedance frequency. At the 2% in 50 year exceedance frequency *VSI*- and *PGV*-based demands are 0.41 m and 0.48 m, respectively. In other words, using *PGV* will over predict the seismic demand on the pile (for a given exceedance rate) by approximately 20% compared to using *VSI*. Investigation of Figure 13-10a in terms of rate of exceedance for a given level of demand indicates that for a pile head displacement of 0.2 m the rates of exceedance are $2.3\text{x}10^{-3}$ and $1.8\text{x}10^{-3}$ using *PGV* and *VSI*, respectively. At a pile head displacement of 0.5 m the rates of exceedance are $3.2\text{x}10^{-4}$ and $2.3\text{x}10^{-4}$ using *PGV* and *VSI*, respectively. Thus at pile head displacements of 0.2 and 0.5 m using *VSI* gives a reduction in the rate of exceedance by 27% and 37%, respectively compared to using *PGV*. Figure 13-10b illustrates the hazard curves for the peak pile curvature and hence damage to the pile using the same two ground motion IMs as above. Superimposed on the figure is the pile curvature levels at which the pile is deemed to exceed yield and ultimate limit states. Thus, it has been assumed that the limit states are a binary function of the demand (in reality the limit states will have some distribution due to cyclic loading effects, material randomness, axial load etc.). For the yielding limit state the annual frequencies of exceedance are 0.025, 0.028 for *VSI*-, and *PGV*-based scaling, respectively. That is, for this hypothetical high seismic region the return period for pile yielding ranges from 36 to 40 years depending on the IM used. For the ultimate limit state the effects of uncertainty are more pronounced with frequencies of $3.5\text{x}10^{-4}$ and $4.5\text{x}10^{-4}$ for the *VSI*- and *PGV*-based scaling, respectively. Thus the return periods for pile failure range from approximately 2200 to 2860 years (a 30% difference). As previously mentioned, the strong correlation between pile curvature and pile head displacement illustrates that *VSI* is an efficient predictor of both the pile head displacement and the pile curvature or damage to the pile.

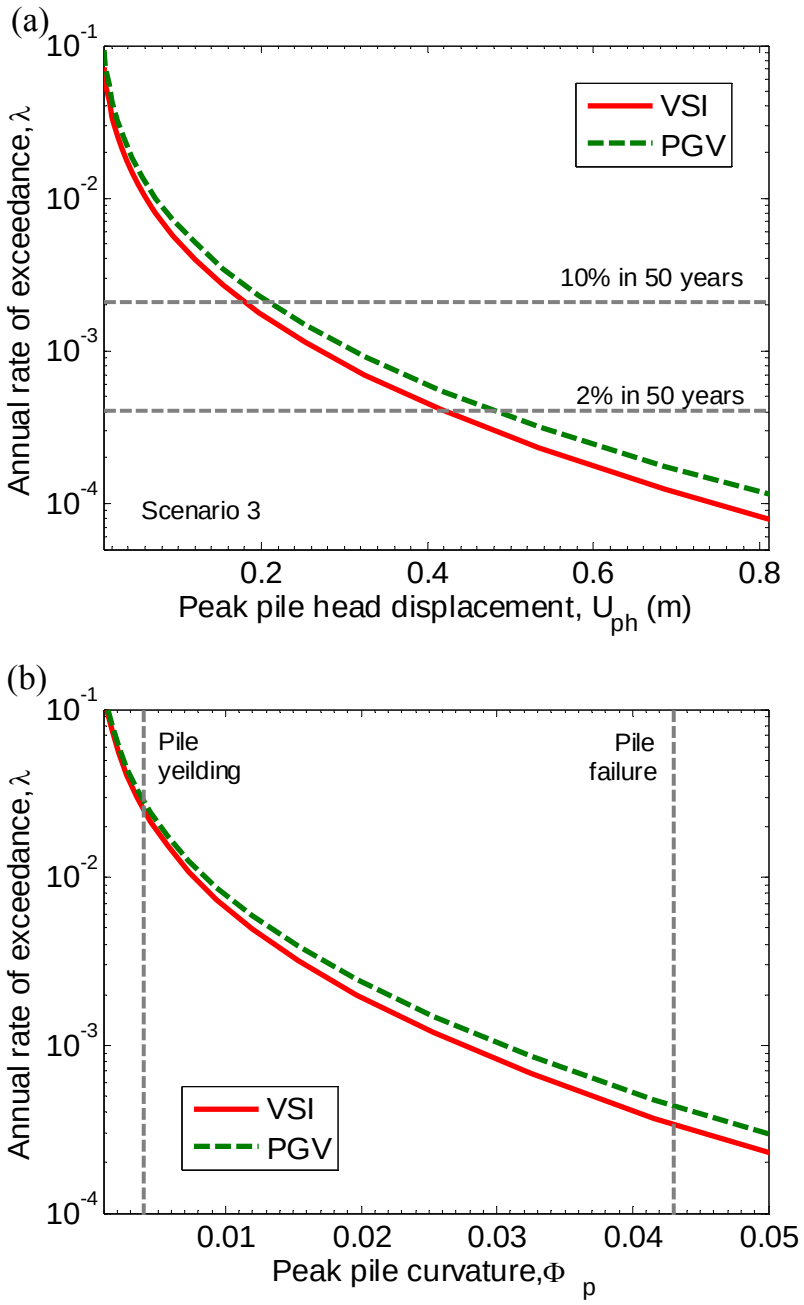


Figure 13-10: Hazard curves for: (a) peak pile head displacement; and (b) peak pile curvature using *VSI* and *PGV* as IMs (scenario 3).

13.7 Intensity measures for liquefiable soils

Thus far the performance of the various intensity measures has been presented with respect to the seismic response analysis for the case of non-liquefiable soils. Seismic response analyses using the same six scenarios presented in Figure 13-2b were also conducted where the effects of excess pore water pressures and eventual liquefaction were explicitly modelled.

As many of the previous comments made regarding IMs for the non-liquefiable soil

scenarios are also applicable to the scenarios using liquefiable soils, then attention here is primarily limited to the discussion of the efficiency of the various IMs. Comparison of the relationship between peak pile curvature and peak pile head displacement from the analyses with liquefiable soils revealed similar trends as that in non-liquefiable soils and therefore peak pile head displacement was again used as the EDP.

13.7.1 Efficiency of IMs for predicting U_{ph} in liquefiable soils

Figure 13-11 illustrates the dispersion in the prediction of the (maximum) pile head displacements as a function of the various IMs for the first four scenarios listed in Figure 13-2b, for simulations with liquefiable soils. Note that several dispersion results were above 0.8 (notably the significant duration, D_s) and are thus not shown in the plot. The variation in the dispersion for a single IM over the four different scenarios illustrates that the specific model configuration does have an effect on the efficiency of the IM. In particular, the effect of the SPT blow count for the top soil layer (either $N^{(1)} = 10$ or 20) appears to be significant for the efficiency of some IMs. This result is distinctly different from that of piles embedded in non-liquefiable soils (Figure 13-6). As was observed in Figure 13-6, Figure 13-11 shows that pile head displacements in liquefiable soils are most efficiently predicted (i.e. show smallest uncertainty) using velocity-based IMs, such as *PGV*, RMS_v, *VSI*, and *SMV*. Peak ground acceleration, *PGA*, and peak ground displacement, *PGD*, provide poor prediction of the response over all four scenarios. Arias Intensity, I_a, which is a measure often used in liquefaction evaluation [29] was found to correlate poorly with pile head displacement, with dispersions above 0.35 for all four scenarios. Spectral acceleration quantities (IM numbers 16-20) tend to predict the pile response in S2 and S4 (where the top layer of soil has $N^{(1)} = 20$) with moderate efficiency (dispersions of approximately 0.35), but their prediction for S1 and S3 (where the top layer of soil has $N^{(1)} = 10$) is very poor. This is due to the larger non-linearity and liquefaction observed in the analyses with the $N^{(1)} = 10$ soil layer compared to the $N^{(1)} = 20$ layer.

Figure 13-11 illustrates that *VSI* is the most efficient IM for predicting the seismic response of piles embedded in liquefiable soils, which was also the most efficient IM for pile response in non-liquefiable soils (Figure 13-6). To illustrate this further, Table 13-3 compares the efficiency of *VSI* in predicting pile head displacements for all six scenarios, for liquefiable and non-liquefiable soil. Three features are apparent upon inspection of Table 13-3: (i) prediction of pile head displacement in liquefying soils is less uncertain than in non-liquefying

soils; (ii) *VSI* is more efficient at predicting peak pile response when the period of the superstructure is T=1.8s (S1 vs S3; S2 vs S4); and (iii) prediction of the pile response when inertial effects are absent (i.e. S5 and S6) is not noticeably improved as compared to the case when both kinematic and inertial effects are present. With respect to the first point, it is important to recall that the analyses with liquefiable soils were conducted up to 0.5g *PGA*, while those for non-liquefiable soils were conducted up to 1.0g *PGA*.

Table 13-3: Efficiency of VSI in predicting peak pile head displacements.

Scenario	Dispersion, $\beta_{\ln U_{ph}\|VSI}$	
	Liquefiable soil	Non-liquefiable soil
S1	0.233	0.255
S2	0.194	0.242
S3	0.261	0.304
S4	0.253	0.288
S5	0.252	0.29
S6	0.265	0.307

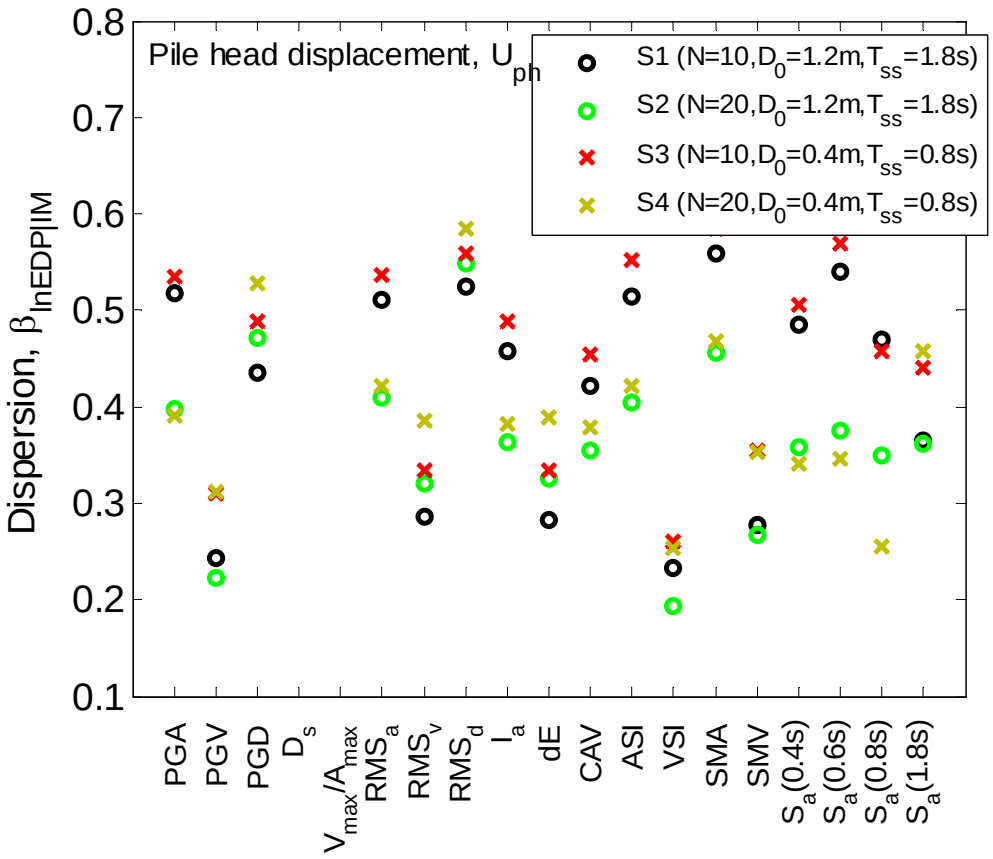


Figure 13-11: Efficiency of various IMs in predicting pile head displacement (liquefiable soils).

13.8 Effects of soil-pile interaction on efficiency

13.8.1 Prediction of ground displacement for liquefiable and non-liquefiable soils

Figure 13-12a illustrates the efficiency of various IMs in predicting the maximum free-field ground displacement, U_g, for the two different soil profiles considered in this study, with a top layer of $N^{(1)} = 10$ or $N^{(1)} = 20$, respectively. Over all of the different IMs considered in the analyses, the free-field soil response is predicted more efficiently for the soil profile having a top layer of $N^{(1)} = 20$ than for a top layer of $N^{(1)} = 10$. This observation is different from the prediction of free-field displacements for non-liquefiable soils shown in Figure 13-12b, where the efficiency is practically independent of the stress-strain characteristics of the soil. This result is due to the pronounced effects of liquefaction on the stress-strain response of the soil with $N^{(1)} = 10$ relative to that with $N^{(1)} = 20$ blow counts. Note that even though liquefaction was induced in the analyses with the $N^{(1)} = 20$ top soil layer, the peak shear strains were significantly less than those developed in the models within the $N^{(1)} = 10$ top soil layer. The reason for the slightly smaller dispersions in free-field soil response for liquefiable soils with $N^{(1)} = 20$ compared to those of non-liquefiable soils, is due to the smaller range of ground motion intensity used in the liquefaction analyses (0.05-0.5g *PGA* and 0.1-1.0g *PGA* for liquefiable and non-liquefiable analyses, respectively).

13.8.2 Efficiency of predicting pile head displacement: effect of relative pile stiffness

Figure 13-13 illustrates the relationship between peak pile head displacement and peak free-field displacement for all of the analyses used in S5 and S6, i.e. those without inertial effects from the superstructure. Figures 13-13a and 13-13b illustrate that during small ground shaking the magnitude of the free-field and pile head displacements are similar for both the 1.2m (S5) and 0.4m (S6) diameter pile. However, as the free-field ground displacements increase (due to an increase in the level of ground motion intensity), significant liquefaction occurs in the upper layer of the soil deposit, and the 1.2m diameter pile exhibits a *relatively*

stiff behaviour, with pile displacements significantly less that that of the free-field soil (i.e. the pile is stiff enough to resist the ground movement). Figure 13-13b illustrates, however, that for the 0.4m diameter pile the relationship between pile head displacement and free-field displacement is almost one-to-one for all ranges of ground motion intensity, and the pile even 'over-shoots' the ground displacement at large intensity levels. This response is typical for a *relatively flexible pile* in which the pile generally follows the ground movement.

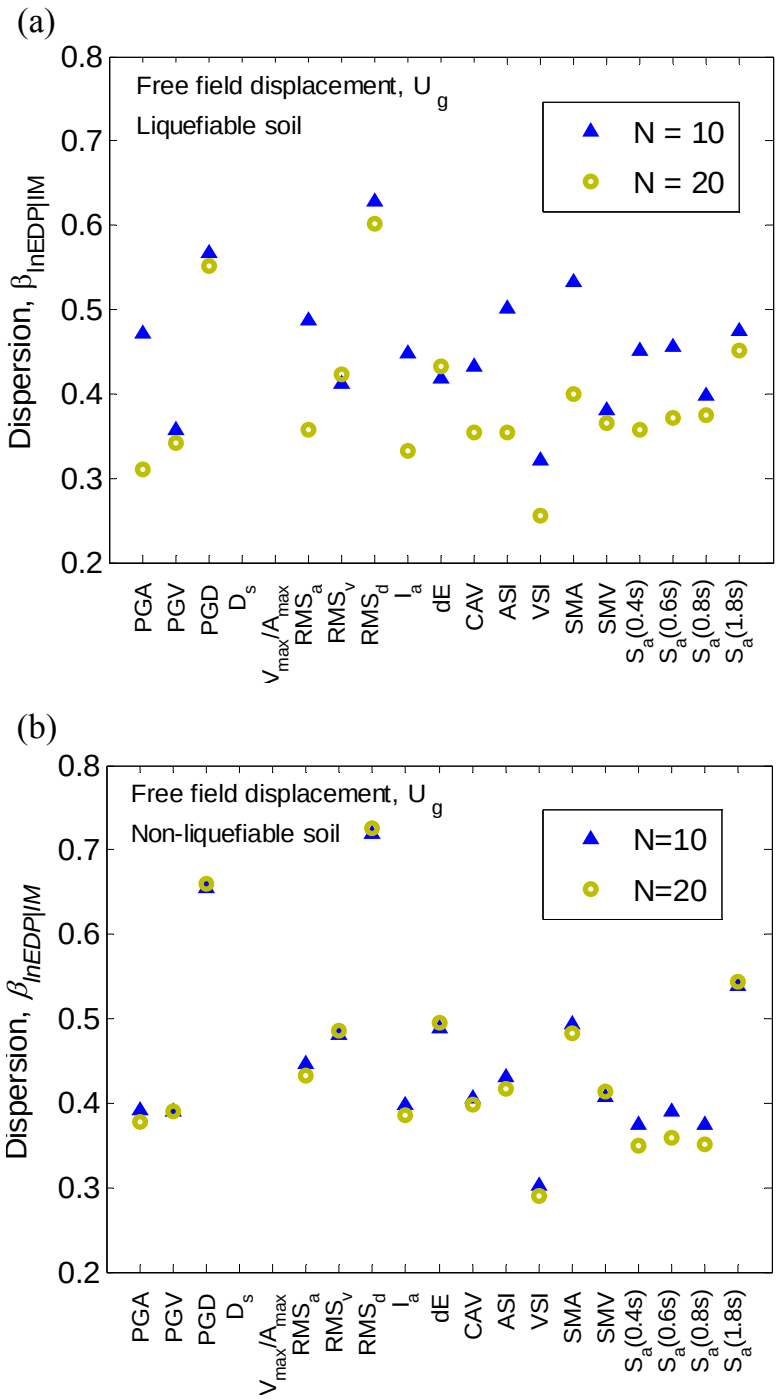


Figure 13-12: Efficiency of various IMs in predicting free-field displacements for the two different upper layer soil types considered: (a) liquefiable soils; (b) non-liquefiable soils.

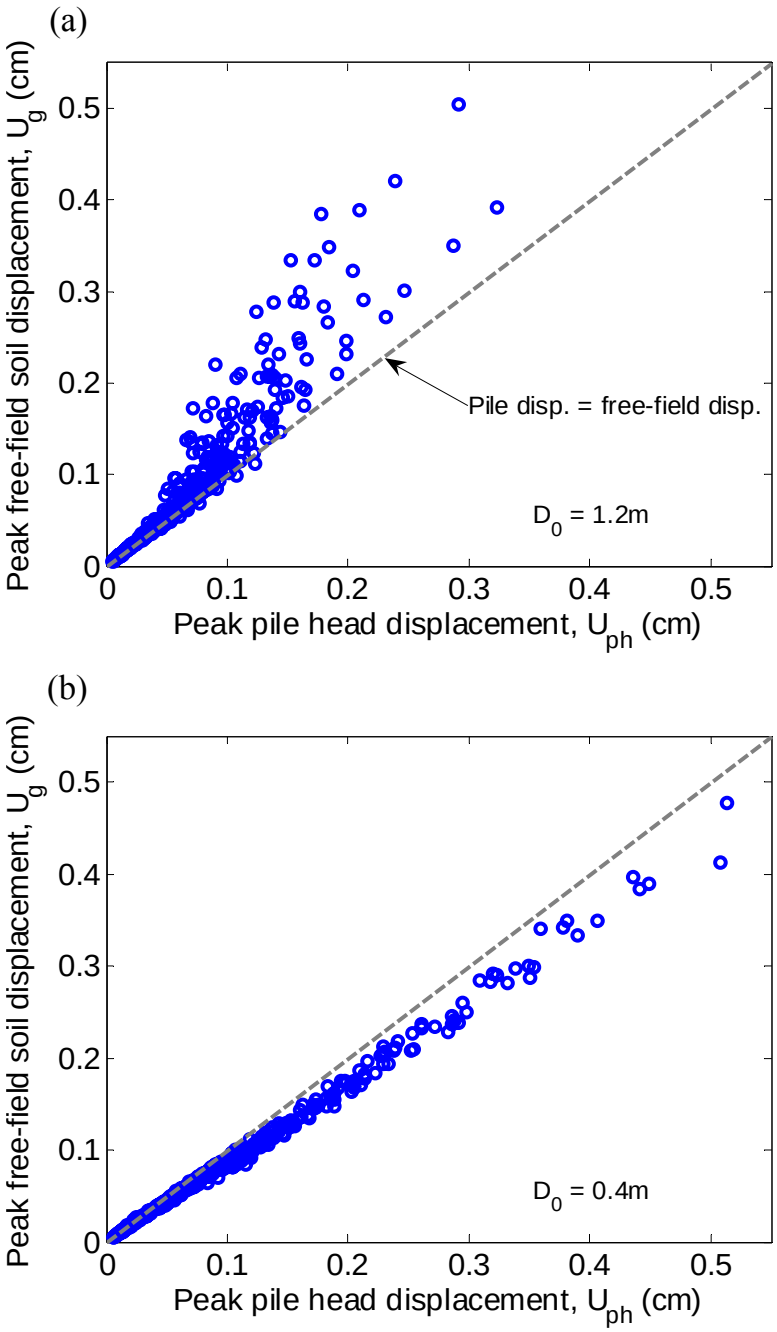


Figure 13-13: Comparison of relationships between free-field soil and pile head displacements for: (a) 1.2m diameter pile; and (b) 0.4m diameter pile (liquefiable soils).

Figure 13-14 illustrates the efficiency in predicting the pile head displacements for S5 and S6 as well as the efficiency in the prediction of the maximum free-field ground displacements. Recall that both S5 and S6 have a top soil layer of $N^{(1)} = 10$ and no superstructure, with the only difference being the diameter of the pile (S5 has $D_0 = 1.2$m and S6 has $D_0 = 0.4$m). Thus, Figure 13-14 allows direct consideration of the effects of the relative pile stiffness on the efficiency of the different IMs, without influence of the

superstructure. Several features of Figure 13-14 are worthy of note: (i) Over the entire range of the examined IMs the efficiency in predicting pile head displacements and free-field ground displacements for the pile with D_o=0.4m (S6) is almost identical, indicating that this pile exhibited a flexible-pile behaviour and moved together with the free-field soil; (ii) the efficiency of predicting the pile head displacements for the stiffer pile with D_o=1.2m (S5) is greater than that for S6 across all of the IMs considered here; and (iii) velocity-based IMs (namely *VSI* and *PGV*) were the most efficient in predicting the free-field ground displacements and thus the imposed kinematic loads on the pile foundation. It is worth noting that the removal of the inertial effects, on average, did not improve the efficiency of the IMs at predicting the peak pile head displacement, as evident from a comparison of Figure 13-14 with S1 and S2 in Figure 13-11.

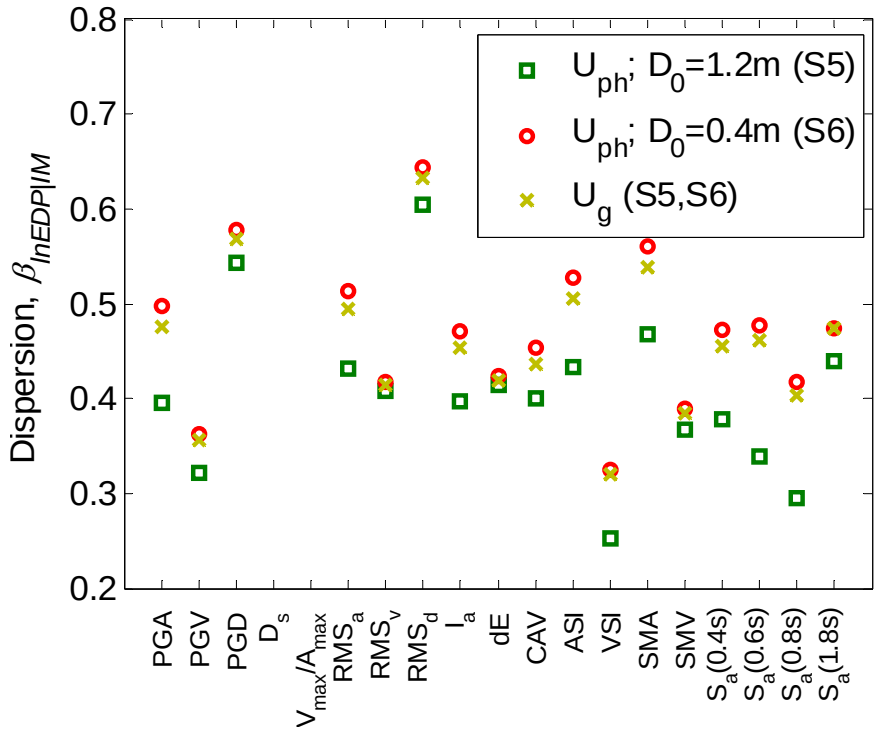


Figure 13-14: Efficiency of various IMs in predicting pile head displacement and free-field soil displacement: illustration of stiff-flexible pile behaviour (liquefiable soils).

13.9 Conclusions

Ground motion intensity measures (IM's) for the evaluation of the seismic performance of pile foundations within the Pacific Earthquake Engineering Research (PEER) Centre framework has been investigated in this study. The seismic response of pile foundations

during strong ground motion is complex and depends on both kinematic effects from the imposed soil displacements and inertial effects from the vibration of the superstructure. A simplified model of a soil-pile-structure system was considered in this study to capture the salient features of the pile response during ground motion excitation. 40 ground motion records were scaled to peak accelerations ranging from 0.1g to 1.0g resulting in 400 seismic analyses for a given scenario. Various soil-pile-structure scenarios were considered allowing for the effects of various soil, pile and superstructure characteristics on the dynamic response of the system. In total, twelve different scenarios were considered; six for piles in non-liquefiable soils and six for piles in liquefiable soils. It was observed that the peak lateral pile head displacement correlated well with the peak curvature of the pile, particularly when kinematic effects dominate the response of the pile foundation. Hence, the peak pile head displacement can be used as an EDP for piles and is an approximate measure for the damage to the pile. The relationship between the peak pile head displacement and the peak pile curvature is not unique, but rather it is a function of the properties of the soil-pile-structure system.

Over all twelve scenarios considered it was found that velocity spectrum intensity (*VSI*) (also known as Housner spectrum intensity, *SI*) predicted the peak pile head displacement with the lowest uncertainty (highest efficiency). This was primarily due to the capability of *VSI* in predicting peak free-field soil displacements which cause kinematic-induced loads on the pile foundation. In addition to having the highest efficiency in the response prediction, *VSI* was found to be sufficient with respect to magnitude, source distance and epsilon, both in absolute terms, and relative to the performance of other IMs e.g. peak ground acceleration and peak ground velocity. *VSI* was also found to be a predictable intensity measure, and it was illustrated how its use can reduce the peak pile head displacement demand hazard compared to other measures of ground motion intensity. The above conclusions apply both to piles in non-liquefiable and liquefiable soils.

The effects of the deformation mechanism of the soil-pile system were also found to be significant with the pile response being more efficiently predicted when *stiff pile* behaviour, as opposed to *flexible pile* behaviour was exhibited for pile response in liquefiable soils.

Despite being based on single-pile models, this study provides clear indications of the efficacy of various IMs in predicting the seismic response of pile foundation within the PBEE framework. Further studies are underway to quantify the effectiveness of the proposed EDPs and IMs for pile groups and capture specific cross-interaction effects on the response of the soil-pile-structure system.

13.10 Acknowledgements

Financial support of the first author by the New Zealand Tertiary Education Commission is greatly appreciated.

13.11 References

[1] Tokimatsu K, Mizuno H, and Kakurai M. Building damage associated with geotechnical problems. Soils and foundations, Special Issue on Geotechnical Aspects of the January 17 1995 Hyogoken-Nambu Earthquake, No. 1 1996: 219–234.

[2] Gazetas G and Mylonakis G. Seismic soil-structure interaction: new evidence and emerging issues, in Geotechnical Earthquake Engineering and Soil Dynamics III (GSP 72,2), 1998, 1119-1174.

[3] Tokimatsu K and Asaka Y. Effects of liquefaction-induced ground displacements on pile performance in the 1995 Hyogoken-Nambu Earthquake. Soils and Foundations 1998: 163-177.

[4] Ishihara K and Cubrinovski M. Case studies of pile foundations undergoing lateral spreading in liquefied deposits, state-of-the-art paper, in 5th International Conference on Case Histories in Geotechnical Engineering, New York, 2004.

[5] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. PEER Center News 2000; 3(2).

[6] Deierlein GG, Krawinkler H, and Cornell CA. A framework for performance-based earthquake engineering, in 7th Pacific Conference on Earthquake Engineering, Christchurch, New Zealand, 2003.

[7] Mehanny SSF and Ayoub AS. Variability in inelastic displacement demands: uncertainty in system parameters versus randomness in ground records. Engineering Structures 2008; 30: 1002-1013., DOI: 10.1016/j.engstruct.2007.06.009.

[8] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35, RMS Program,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[9] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthquake Spectra 2007; 23(2): 357-392, DOI: 10.1193/1.2723158.

[10] Kramer SL and Mitchell RA. Ground motion intensity measures for liquefaction hazard evaluation. Earthquake Spectra 2006; 22(2): 413-438.

[11] Diana-J3. Finite-element program for effective stress analysis of two-phase soil medium. 1997.

[12] Cubrinovski M and Ishihara K. Modelling of Sand Behaviour based on state concept. Soils and Foundations 1998; 28(3): 115-127.

[13] Cubrinovski M and Ishihara K. State Concept and Modified Elastoplasticity for Sand Modelling. Soils and foundations 1998; 38(4): 213-225.

[14] Cubrinovski M, Uzuoka R, Sugita H, Tokimatsu K, Sato M, Ishihara K, Tsukamoto Y, and Kamata T. Prediction of pile response to lateral spreading by 3-D soil-water coupled dynamic analysis: shaking in the direction of ground flow. Soil Dynamics and Earthquake Engineering 2008; 28(6): 421-435, DOI: 10.1016/j.soildyn.2007.10.015.

[15] Seed RB and Idriss IM. Soil moduli and damping factors for dynamic response analysis. Earthquake Engineering Research Center, UC Berkeley, CA, 1970. 48.

[16] Medina RA and Krawinkler H. Seismic demands for nondeteriorating frame structures and their dependence on ground motions. John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA., 2003. 373 pp.

[17] Tothong P and Luco N. Probabilistic Seismic Demand Analysis Using Advanced Ground Motion Intensity Measures. Earthquake Engineering and Structural Dynamics 2007; 36(13): 1837-1860.

[18] Baker JW and Cornell CA. A vector-valued ground motion intensity measure consisting of spectral acceleration and eplison. Earthquake Engineering and Structural Dynamics 2005; 34(10): 1193-1217.

[19] Riddell R. On ground motion intensity measures. Earthquake Spectra 2007; 23(1): 147-173.

[20] Park YJ and Ang AHS. Mechanistic Seismic Damage Model for Reinforced Concrete. Journal of Structural Division, ASCE 1985; 111(4): 722-739.

[21] Porter K, Kennedy R, and Bachman R. Creating Fragility Functions for Performance-based Earthquake Engineering. Earthquake Spectra 2007; 23(2): 471-489.

[22] Cubrinovski M and Ishihara K. Simplified method for analysis of piles undergoing lateral spreading in liquefied soils. Soils and Foundations 2004; 44(25): 119-133.

[23] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[24] Housner GW. Spectrum intensities of strong-motion earthquakes, in Symposium on earthquakes and blast effects on structures, Los Angeles, CA, 1952.

[25] Aslani H and Miranda E. Probability-based Seismic Response Analysis. Engineering Structures 2005; 27(8): 1151-1163.

[26] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges. Engineering Structures 2007; 29(10): 2662-2672, DOI: 10.1016/j.engstrut.2006.12.015.

[27] Danciu L and Tselentis GA. Engineering Ground motion Attenuation Relationships for Greece. Bulletin of the Seismological Society of America 2007; 97(1B): 162-183, DOI: 10.1785/0120040087.

[28] Bradley BA, Cubrinovski M, MacRae GA, and Dhakal RP. Ground motion prediction equation for spectrum intensity from spectral acceleration relationships. Bulletin of the Seismological Society of America 2009; 99(1): 277-285, DOI: doi: 10.1785/0120080044.

[29] Kayen RE and Mitchell JK. Assessment of liquefaction potential during earthquakes by Arias intensity. Journal of Geotechnical and Geoenvironmental Engineering 1997; 123(12): 1162–1174.

14.Probabilistic Seismic Performance and Loss Assessment of a Bridge-Foundation-Soil System

Bradley BA, Cubrinovski M, Dhakal RP, MacRae GA. Probabilistic seismic performance and loss assessment of a bridge-foundation-soil system. *Soil Dynamics and Earthquake Engineering* (submitted).

14.1 Abstract

This paper presents the probabilistic seismic performance and loss assessment of an actual bridge-foundation-soil system, the Fitzgerald Avenue twin bridges in Christchurch, New Zealand. A two-dimensional finite element model of the longitudinal direction of the system is modelled using advanced soil and structural constitutive models. Ground motions at multiple levels of intensity are selected based on the seismic hazard deaggregation at the site. Based on rigorous examination of several deterministic analyses, engineering demand parameters (EDP's) which capture the global and local demand and consequent damage to the bridge and foundation are determined. A probabilistic seismic loss assessment of the structure considering both direct repair and loss of functionality consequences was performed to holistically assess the seismic risk of the system.

It was found that the non-horizontal stratification of the soils, liquefaction, and soil-structure interaction had pronounced effects on the seismic demand distribution of the bridge components, of which the north abutment piles and central pier were critical in the systems seismic performance. The consequences due to loss of functionality of the bridge during repair were significantly larger than the direct repair costs, with over a 2% in 50 year probability of the total loss exceeding twice the book-value of the structure.

14.2 Introduction

Methods for assessment of the seismic performance of soil-structure systems have evolved significantly in the past two decades. This evolution has involved further improvement of simplified design-oriented approaches, and also development of more robust, and complex, analysis procedures. In addition to the development in methods of analysis, attention has shifted from the implicit assessment of seismic performance via seismic response analysis, to an explicit consideration of seismic performance based on the societal and economic consequences of seismic response and associated damage.

Consideration of the seismic response of soil-structure systems is complicated by the complexity of the ground motion excitation and the non-linear dynamic response of soil-structure systems. In addition to this complexity, the seismic response of soil-structure systems is burdened by a significant amount of uncertainty. Such uncertainty arises due to the uncertain nature of the future ground motions at the site, as well as the lack of knowledge of the properties governing the response of the soil-structure system. In addition to the ground motion and seismic response uncertainties, there are also uncertainties associated with the levels of damage to the structure and the corresponding consequences in terms of direct repair costs, loss of functionality, and human injuries.

Recent efforts [1-3], predominantly following the Pacific Earthquake Engineering Research (PEER) Centre framework formula [4], have focused on performance-based methodologies which allow the computation of seismic performance measures encompassing the direct and indirect consequences associated with the seismic response of engineered facilities, as well as addressing the significant aforementioned uncertainties in the seismic assessment problem.

Such performance and loss assessment methodologies have been primarily applied to assess the direct repair loss to structural systems such as office buildings [1-3, 5], with less attention to date devoted to the consideration of lifelines such as bridge structures, particularly those in which soil liquefaction can result in significant ground motion modification and demand to the structure and foundations. Furthermore, due to their increased complexity, indirect consequences due to loss of functionality have also not commonly received attention.

The focus of this paper is the probabilistic seismic performance and loss assessment of a two-span bridge structure supported on pile foundations which are founded in liquefiable soils. Firstly, the structure, site conditions, and computational model of the bridge-

foundation-soil system are discussed. An overview of the seismic response of the system for a single ground motion is discussed to elucidate the predominant deformation mechanisms of the system and to identify the engineering demand parameters (EDP's) to use in the probabilistic seismic demand and loss assessments. Ground motions are selected in accordance with the seismic hazard deaggregation for various intensity levels, and the results of the seismic response analyses are used to perform probabilistic seismic demand and loss assessments of the system considering repair cost and loss of functionality consequences.

14.3 Case study: Fitzgerald Avenue bridges

14.3.1 Details of the structure

The Fitzgerald Avenue twin bridges are located to the north-west of central Christchurch, New Zealand. Each of the two-span bridges is 30 m long, 12.1 m wide and 3.2 m high (Figure 14-1a). The 15 m bridge deck spans consist of 21 prestressed concrete I-girders and cast-in-place concrete slabs. The bridge superstructure is supported on two seat abutments and one central pier (Figure 14-1b). The abutments and pier are 2.5 m high and are supported on pile foundations consisting of 8 x 0.3 m diameter piles. All piles have continuous moment connections at the pile cap. At both abutments the bridge deck is seated on a 10 mm bearing pad as illustrated in Figure 14-1c.

The Fitzgerald Avenue bridges are a key link in Christchurch's transportation network carrying 38,000 vehicles daily. Because of their location in the transportation network, the Fitzgerald Avenue bridges have been designated by the Christchurch City Council as a key lifeline for post-earthquake services. A recent assessment of the existing bridge structure [6] recommended the installation of two additional bored piles at each of the abutments and central pier to a depth of 25 m. These two additional piles on each side of the central pier are 1.5 m in diameter, while those at the abutments are 1.2 m in diameter.

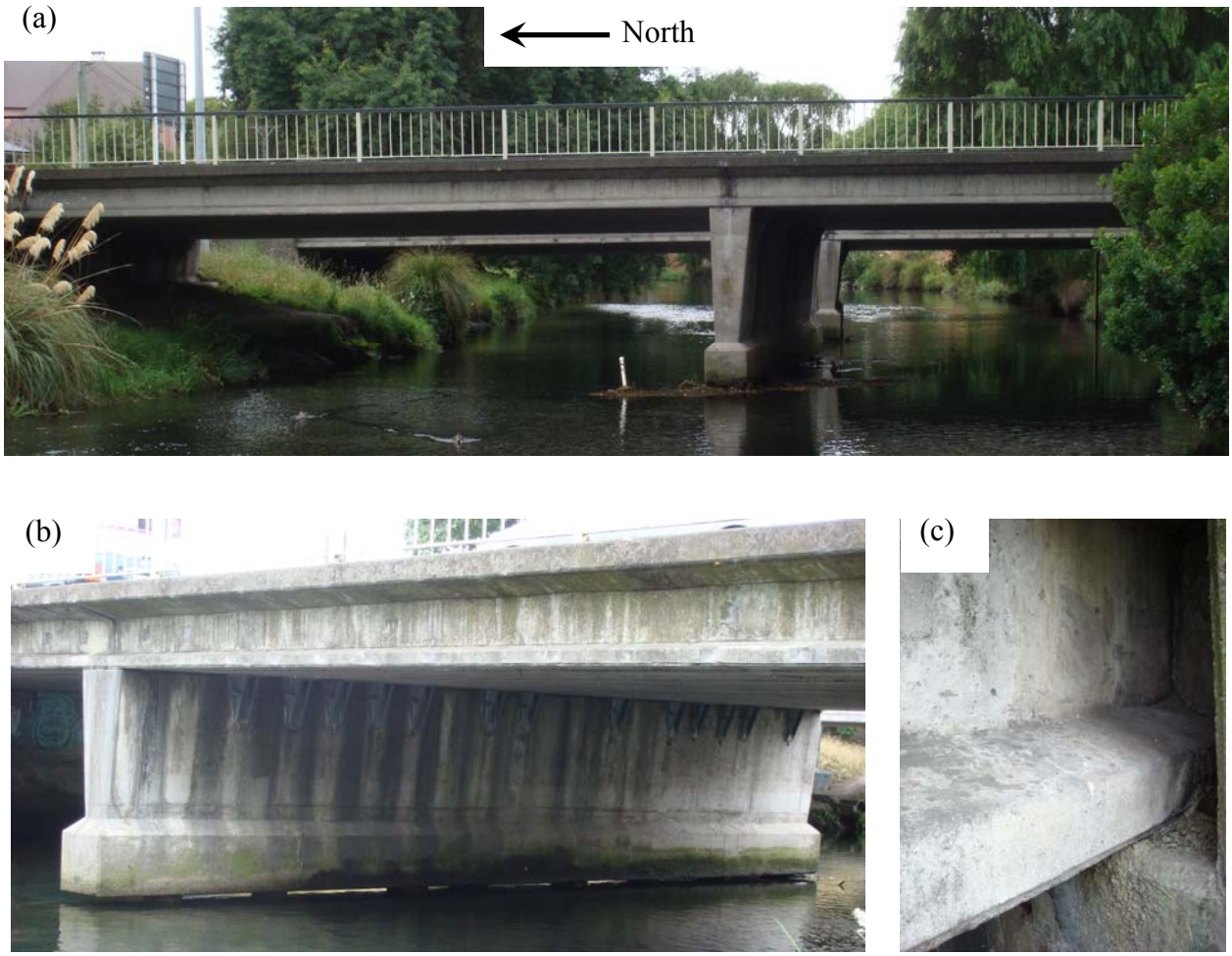


Figure 14-1: The Fitzgerald Avenue twin bridges: (a) elevation of the west bridge; (b) central pier and pile cap; and (c) seating connection of bridge deck on abutments. North is indicated in the direction right to left.

14.3.2 Site conditions

Previous site investigations conducted to confirm ground conditions and assess material properties and liquefaction potential include: standard penetration tests (SPT's); cone penetration tests (CPT's) with direct push Dual Tubes (DT's); and the installation of piezometers. Based on these field investigations the generic soil profile for the longitudinal axis of the bridge given in Figure 14-2 was developed. The soil profile consists of four distinct layers. The shallowest two horizontal layers have thicknesses of 4.5 m and 6.5 m, and normalised SPT blowcounts of $N_1 = 10$ and $N_1 = 15$, respectively. Below these two layers, the profile deviates from a simple horizontal layering, with a weaker layer of 6.5 m depth and SPT blowcount of $N_1 = 10$ on the left hand side of the model. Below 17.5 m on the left hand side of the model, and up to 11m depth on the right hand side of the model is a significantly stiffer layer of $N_1 = 30$. Both the $N_1 = 10$ and $N_1 = 15$ layers are highly susceptible to

liquefaction, while the $N_1 = 30$ base layer was deemed to be non-liquefiable. Behind the abutments, gravel backfills extend at an angle of 30 degrees to the ground surface.

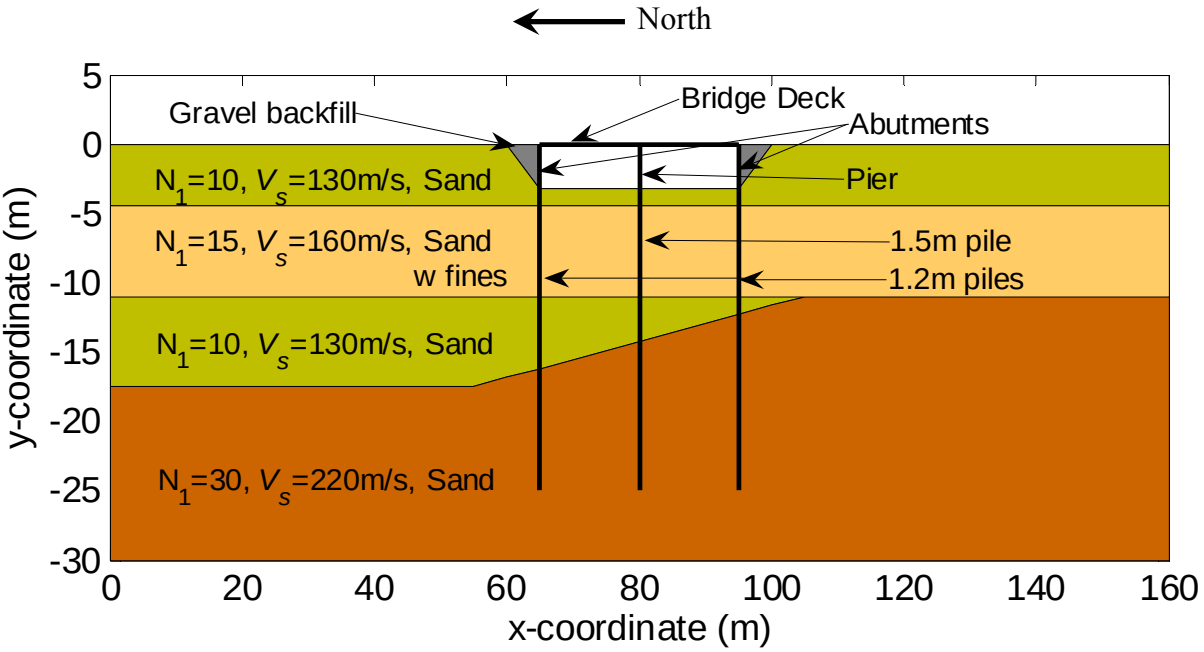


Figure 14-2: Schematic illustration of computational model. Soil, pile and superstructure properties are given in Table 14-1.

Table 14-1: Pile and superstructure model properties.

Structural element	M_C, M_Y, M_U (MN-m)	ϕ_C , ϕ_Y , ϕ_U (1/m)	EI , M_F[1] (MN-m^2) (MN-m)
1.2m diameter pile	1.25, 4.22, 6.74	0.0003,0.0018, 0.009	4.1x10^3 , 7.70
1.5m diameter pile	2.75, 7.50, 12.0	0.0005, 0.0015, 0.009	9.7x10^3 , 14.0
Bridge pier	1.58, 4.91, 6.50	0.00055, 0.0036, 0.01	3.6x10^3 , 7.91
Bridge deck[2]	-	-	3.99x10^3

[1] EI, M_F values are those used in the hyperbolic moment-curvature model.
[2] Bridge deck modeled as linear elastic.

14.4 Computational model

A non-linear plane-strain finite element model of the longitudinal direction of the bridge-foundation-soil system was constructed using a seismic effective stress method of analysis [7]. While the seismic response of the bridge-pile-soil system is clearly a 3-dimensional problem, only the analyses of the longitudinal direction are discussed herein. Details of the seismic effective stress analyses of the transverse direction of the bridge system are presented in Bowen and Cubrinovski [8] and Cubrinovski and Bradley [9].

Because of symmetry, the out-of-plane width of the longitudinal plane-strain model was taken to be half of the bridge width (6.05 m). That is, half of the bridge deck, abutments and

piers were considered, as well as the same dimension for the soil thickness. Therefore, in the computational model, each abutment and the central pier are supported by a single 1.2 m and 1.5 m diameter pile, respectively. The 0.3 m diameter piles which supported the structure before the installation of the 1.2 m and 1.5 m piles provide negligible contribution and were not considered in the computational model.

The soil was modelled using an elastic-plastic constitutive model (S-D model) particularly tailored for modelling liquefaction problems [10, 11]. The model combines two fundamental sand modelling concepts. The first is the state concept [12], where the sand behaviour is characterised based on its density and confining stress. The second is a modified elastic-plastic formulation with continuous yielding and hypoplasticity. The S-D model has been extensively verified through rigorous simulations of down-hole array records at liquefaction sites, seismic centrifuge tests, large-scale shake-table tests on pile foundations and case histories of damaged piles in the 1995 Kobe earthquake (e.g. [13] and references therein). Further details on the computation of the constitutive model parameters used in the analysis is given in Bowen and Cubrinovski [8] and Cubrinovski and Ishihara [10, 11].

The bridge abutments, central pier and pile foundations were modelled using non-linear beam elements. The moment-curvature response was parameterized by a hyperbolic curve, with the initial stiffness, EI, and peak moment, M_F, chosen to match the moment curvature relationship of the pile (See Bowen and Cubrinovski [8] for details). The unloading/reloading path for the moment-curvature relationship is based on the Masing rule [14], and no strength degradation was considered due to limitations of the constitutive model. The bridge superstructure was modelled as linear elastic because of its significantly higher axial stiffness compared to the lateral stiffness of the abutments/piers and its higher flexural and shear strength.

A static analysis of the model was performed in order to determine the initial stress distribution in the model. In particular, a correct distribution of shear stresses near the abutments is critical for modelling the driving stresses for lateral spreading of soil toward the river channel. In addition to hysteretic damping occurring as a result of the inelastic constitutive models, Rayleigh damping was used to provide enhanced numerical stability with parameters $\alpha = 0$ and $\beta = 0.005$.

14.5 Seismic hazard and ground motions

The seismic hazard due to earthquake-induced ground motion is determined using

probabilistic seismic hazard analysis (PSHA) [15]. In order to obtain the seismic hazard curve it is first necessary to specify which ground motion intensity measure (IM) is to be used. In this study, *PGA* is used as the IM, both for its historical use and because it, and spectral accelerations at various periods, are the only IM's for which seismic hazard curves are publicly available for this site. Recent studies [16] have shown however that velocity-based IM's (e.g. peak ground velocity, *PGV*, and spectrum intensity, *SI*) are better IM's for such analyses of structures in liquefiable soils. As advanced ground motion prediction equations for these IM's are available (e.g. Bradley *et al*. [17] and Boore and Atkinson [18]), the development of seismic hazard curves for such IM's should be a future focus in performance-based geotechnical earthquake engineering. As will be seen in the following sections, the choice of *PGA* as the IM leads to significant scatter in the results of the seismic response analyses.

Figure 14-3a illustrates the ground motion hazard at the site of the bridge structure, while Figure 14-3b and Figure 14-3c illustrate the deaggregation and target spectra necessary for ground motion selection [19]. Ground motion selection in accordance with the seismic hazard deaggregation has been shown to be important [20], particularly for inefficient and insufficient IMs such as *PGA*. As noted in Stirling [21], and evident in Figure 14-3b, the seismic hazard is dominated by: (i) $M_W = 5.5\text{-}6.5$ earthquakes at short distances ($R = 15\text{-}30$ km), associated with background seismicity, and (ii) larger $M_W = 6.9\text{-}7.6$ earthquakes on mapped faults ranging from $R = 25\text{-}50$ km.

Ground motions were selected for seismic response analyses at 9 different intensity levels as shown in Figure 14-3a and Table 14-2. For each intensity level, ground motions were selected from the NGA database [22] based on the M_w, R and ε deaggregation from PSHA (e.g. Figure 14-3b). A further constraint of an amplitude scale factor in the range, $SF = 0.6\text{-}1.6$, was used to reduce response bias [23] and help ensure that ground motions with the correct frequency content (i.e. spectral shape) were selected. Because of the larger recurrence intervals and resulting ground motions of earthquakes on mapped faults, deaggregation at large exceedance probabilities was dominated by the background seismicity, while the hazard at small exceedance probabilities was dominated by the mapped faults. This source variation was considered in ground motion selection by, for example: 90% of the selected ground motions being based on scenario type (i) and 10% on scenario type (ii) for the lowest IM level; 80% vs. 20% at IM level two and so on (i.e. at IM level 9 there were 10% of ground motions due to scenario type (i), and 90% of type (ii)).

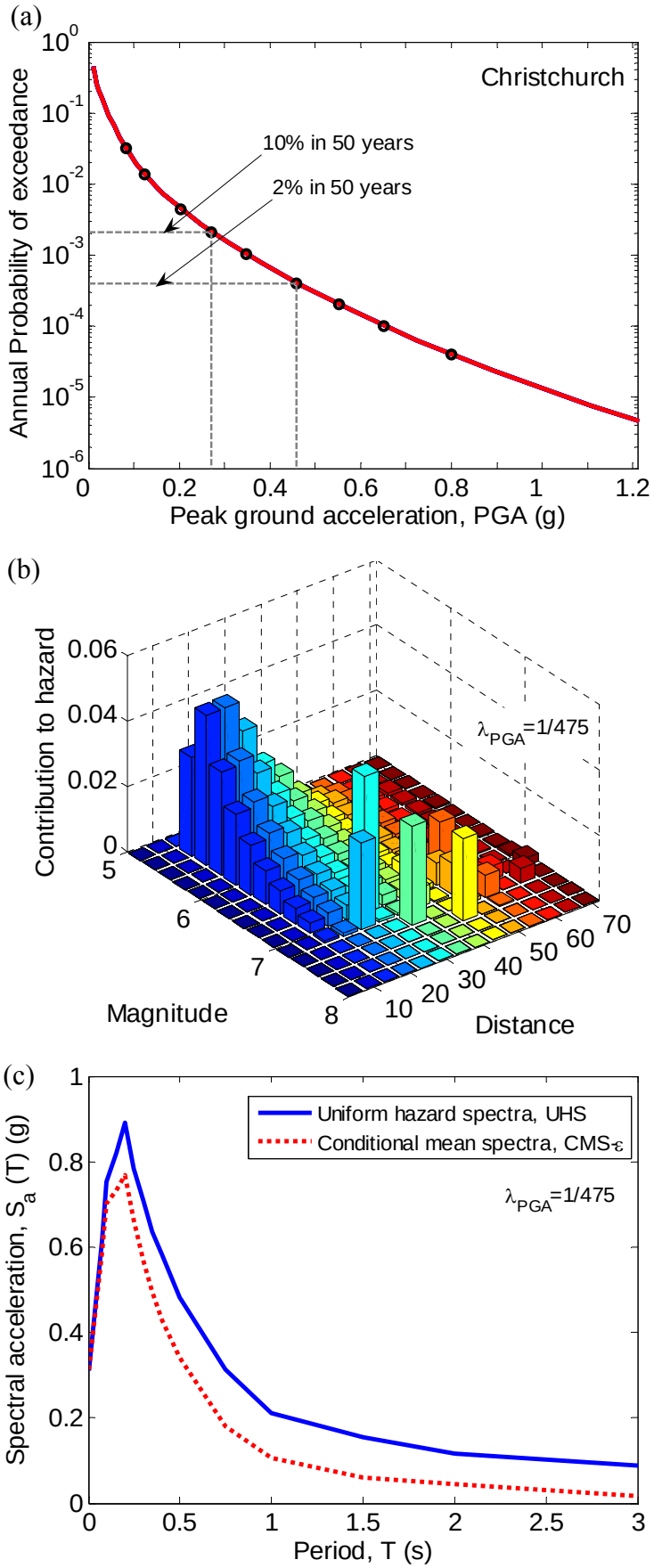


Figure 14-3: Details of the PGA seismic hazard for class C soil in Christchurch: (a) Seismic hazard curve; (b) Deaggregation of the hazard curve for λ_{PGA}=1/475; and (c) Ground motion target spectra for λ_{PGA}=1/475.

Table 14-2: Ground motion intensity levels and their rate of exceedance.

IM level	*PGA*	P(*PGA*>*pga*|50yr)	λ_{PGA}
1	0.0834	0.8	3.22×10^{-2}
2	0.125	0.5	1.39×10^{-2}
3	0.205	0.2	4.46×10^{-3}
4	0.276	0.1	2.11×10^{-3}
5	0.355	0.05	1.03×10^{-3}
6	0.463	0.02	4.04×10^{-4}
7	0.562	0.01	2.01×10^{-4}
8	0.671	0.005	1.00×10^{-4}
9	0.825	0.002	4.00×10^{-5}

Based on the above criteria, the ground motions shown in Table 14-3 were selected and used in the seismic response analyses to follow. For each IM level, ten ground motion records, each containing two orthogonal ground motion components, were selected, giving a total of 180 ground motions for seismic response analyses. Ground motions were linearly (i.e. amplitude) scaled based on the rotation independent geometric mean of the two orthogonal components [24]. Because of the range of scale factors allowed (0.6<SF<1.6) some ground motions were acceptable for two or more IM levels as illustrated in the right-hand column of Table 14-3.

14.6 Deterministic performance assessment

Before conducting the probabilistic seismic response analyses with multiple ground motions and at multiple intensity levels, it was considered important to first rigorously examine the characteristics of the response of the system using detailed seismic effective stress analysis. This is important for understanding the development of excess pore water pressures; ground response; soil-structure interaction; and the predominant deformation mechanisms which control the seismic response of the system. The latter point, in particular, is necessary before conducting probabilistic effective stress analyses since the number of analyses employed in the probabilistic assessment means it is not feasible to examine each analysis in detail, but rather a set of engineering demand parameters (EDP's) are used to indicate the seismic response. Thus an understanding of the deformational mechanism and response of the system is critical in the selection of appropriate EDP's. To that goal, this section discusses the seismic response of the computational model for a single ground motion scaled to an intensity level with a 2% probability of exceedance in 50 years.

Table 14-3: Ground motions used in the seismic effective stress analyses.

ID[1]	Event	Year	Station	M	R (km)	V_{s30} (m/s)	PGA (g)	IM level
15	Kern County	1952	Taft Lincoln School	7.36	38.89	385.4	0.173	3
33	Parkfield	1966	Temblor pre-1969	6.19	16.24	527.9	0.293	4, 5
57	San Fernando	1971	Castaic - Old Ridge Route	6.61	22.63	450.3	0.299	6
68			LA - Hollywood Stor FF	6.61	25.89	316.5	0.210	4
77			Pacoima Dam (left abut)	6.61	3.03	2016.1	1.164	9
154	Coyote Lake	1979	San Juan Bautista, 24 Polk St	5.74	19.70	370.8	0.101	1, 2
189	Imperial Valley-06	1979	SAHOP Casa Flores	6.53	10.79	338.6	0.357	6
231	Mammoth Lakes-01	1980	Long Valley Dam (L Abut)	6.06	15.46	345.4	0.340	5
265	Victoria, Mexico	1980	Cerro Prieto	6.33	14.37	659.6	0.572	7
316	Westmorland	1981	Parachute Test Site	5.90	16.81	348.7	0.219	3, 4
318			Superstition Mtn Camera	5.90	19.50	362.4	0.101	1, 2
359	Coalinga-01	1983	Parkfield - Vineyard Cany 1E	6.36	26.38	338.5	0.182	3
450	Morgan Hill	1984	Corralitos	6.19	23.43	462.2	0.098	1, 2
534	N. Palm Springs	1986	San Jacinto - Soboba	6.06	23.31	370.8	0.231	3, 4
552	Chalfant Valley-02	1986	Lake Crowley - Shehorn Res.	6.19	24.47	338.5	0.123	1, 2
553			Long Valley Dam (Downst)	6.19	21.12	345.4	0.075	1
598	Whittier Narrows-01	1987	Big Tujunga, Angeles Nat F	5.99	28.50	446.0	0.149	1, 2, 3
600			Brea Dam (Downstream)	5.99	23.99	370.8	0.231	4
611			Compton - Castlegate St	5.99	23.37	308.6	0.331	5
625			Inglewood - Union Oil	5.99	25.86	316.0	0.263	3, 4
626			LA - 116th St School	5.99	23.29	301.0	0.341	4, 5
638			LA - N Westmoreland	5.99	21.11	315.1	0.201	3
641			LA - Saturn St	5.99	24.99	308.7	0.123	1, 2
692			Santa Fe Springs - E.Joslin	5.99	18.49	308.6	0.433	5, 6, 7
727	Superstition Hills-02	1987	Superstition Mtn Camera	6.54	6.56	362.4	0.793	8, 9
767	Loma Prieta	1989	Gilroy Array #3	6.93	12.82	349.9	0.462	6, 7, 8
770			Gilroy Array #7	6.93	22.68	333.9	0.312	4, 5
776			Hollister - South & Pine	6.93	27.93	370.8	0.279	5
779			LGPC	6.93	3.88	477.7	0.784	8, 9
810			UCSC Lick Observatory	6.93	18.41	714.0	0.457	7
811			WAHO	6.93	17.47	376.1	0.517	7, 8
828	Cape Mendocino	1992	Petrolia	7.01	8.18	712.8	0.624	8, 9
830			Shelter Cove Airport	7.01	28.78	513.7	0.195	2, 3
838	Landers	1992	Barstow	7.28	34.99	370.8	0.119	1, 2
879			Lucerne	7.28	3.71	684.9	0.721	9
900			Yermo Fire Station	7.28	23.80	353.6	0.223	3, 4
952	Northridge-01	1994	Beverly Hills - 12520 Mulhol	6.69	18.36	545.7	0.510	7, 8
963			Castaic - Old Ridge Route	6.69	20.72	450.3	0.490	5, 6, 7
982			Jensen Filter Plant	6.69	5.43	373.1	0.764	9
983			Jensen Filter Plant Generator	6.69	5.43	525.8	0.765	9
991			LA - Cypress Ave	6.69	30.70	446.0	0.206	4
995			LA - Hollywood Stor FF	6.69	24.03	316.5	0.335	6
998			LA - N Westmoreland	6.69	26.73	315.1	0.370	5, 6
999			LA - Obregon Park	6.69	37.36	349.4	0.467	6, 7
1001			LA - S Grand Ave	6.69	33.99	308.6	0.273	5
1003			LA - Saturn St	6.69	27.01	308.7	0.454	6, 7
1004			LA - Sepulveda VA Hospital	6.69	8.44	380.1	0.803	9
1007			LA - Univ. Hospital	6.69	34.20	376.1	0.349	6
1054			Pardee - SCE	6.69	7.46	345.4	0.505	8
1077			Santa Monica City Hall	6.69	26.45	336.2	0.591	7, 8
1080			Simi Valley - Katherine Rd	6.69	13.42	557.4	0.745	8, 9
1085			Sylmar - Converter Sta East	6.69	5.19	370.5	0.647	8, 9
1643	Sierra Madre	1991	LA - City Terrace	5.61	25.69	365.2	0.102	1, 2
1647			San Marino - SW Academy	5.61	18.74	379.4	0.144	1, 2, 3

[1] As given in the NGA database. http://peer.berkeley.edu/nga/earthquakes.html

14.6.1 Foundation soil response

Figure 14-4 illustrates the scaled acceleration time history and acceleration response spectra for the Loma Prieta Gilroy Array #3 $\phi = 90$ ground motion (ID = 767 in Table 14-3), which was used as the input ground motion in the deterministic seismic response analysis. The geometric mean of this motion (not the individual component) was scaled to the 2% in 50 year exceedance value of 0.463g *PGA* (*SF* = 1.002).

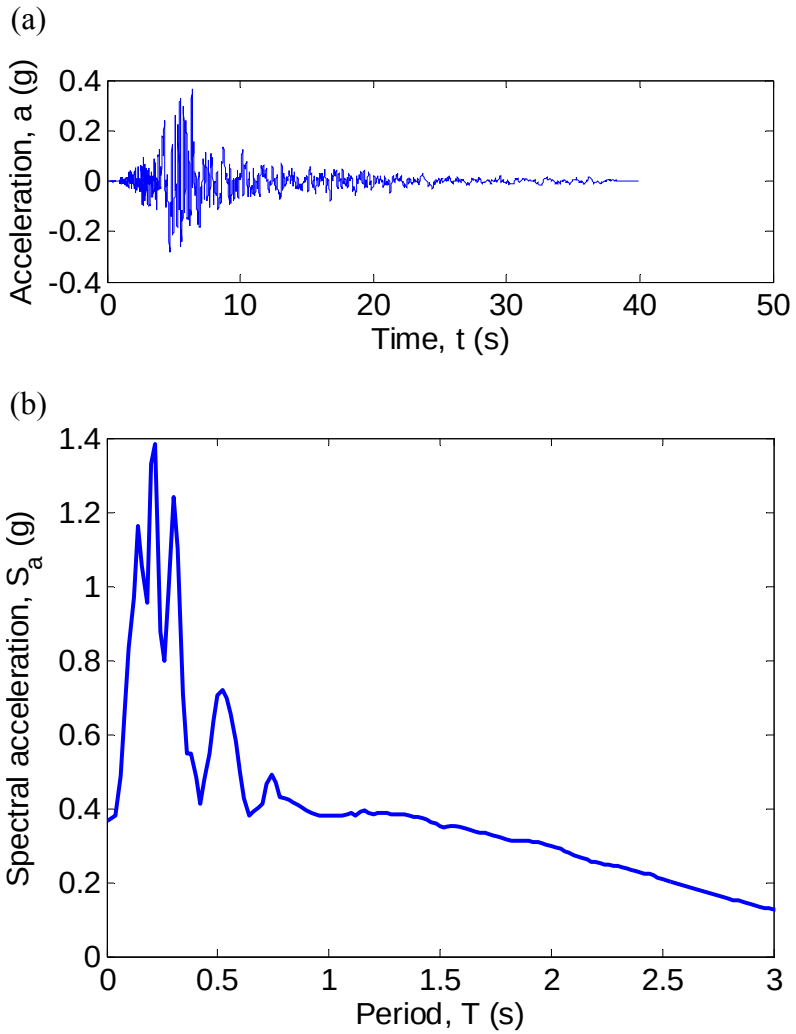


Figure 14-4: (a) acceleration time-history; and (b) acceleration response spectra of the ground motion used in the deterministic performance assessment.

Figure 14-5 illustrates the development of excess pore pressures and eventual liquefaction in the soil surrounding the bridge. It can be seen that pore pressure ratios in the range, *EPWPR* = 0.2-0.5, first develop in the bottom N_1 = 10 layer on the left hand side of the model and at the base of the N_1 = 15 layer on the right hand side of the model. The bottom N_1 = 10 layer has almost entirely liquefied by 6.0 seconds. As shaking progresses with time,

pore water pressures continue to increase in the right hand side of the N_1 = 15 layer, and the re-distribution of excess pore pressures causes spreading of liquefaction to shallower depths (predominantly on the left hand side of the model).

Figure 14-6a illustrates excess pore water pressure ratios 45 m to the left of the bridge. The three depths of z = 6.15, 14.75, and 19.75 m are located in the N_1 = 15, 10, and 30 layers, respectively. In agreement with Figure 14-5, it can be seen that complete liquefaction of the N_1 = 10 (i.e. z = -14.75m) layer by t = 7.0 s. Liquefaction of the bottom N_1 = 10 layer also reduces the ground motion intensity in the above soil layers, preventing liquefaction from eventuating at z = 6.15m. Figure 14-6b illustrates the shear stress-strain response of the soil at z = -14.75m, during which liquefaction and peak shear strains of up to 2.5% were computed.

14.6.2 Bridge and pile response

Figure 14-7a illustrates the displacement time histories at the three footings of the bridge, and the north and south free-field response (all at a depth of z = -3.2 m). In the first 7.0 s, it is apparent that the displacement in the north free-field is larger than the south free-field and footing displacements, which are essentially identical. After t = 7.0 s, relative displacements between the three footings are apparent due to significant liquefaction occurring in the surrounding soils. It is also apparent in Figure 14-7a that the displacement histories of the footings are not completely in-phase with the free-field response (both north and south). Figure 14-7b provides a comparison of the acceleration histories at the north free-field (z = 0 m), central pile cap, and at 27.5 m depth, near the base of the model. It can be seen that the stiffening effect of the pile foundations allows waves of significantly higher amplitude and frequency to propagate to the central pier cap than to the free-field surface where significant soil liquefaction occurs. This stiffening effect is the reason for the aforementioned out-of-phasing and smaller amplitude of the footing displacements in Figure 14-7a compared to that in the north free-field.

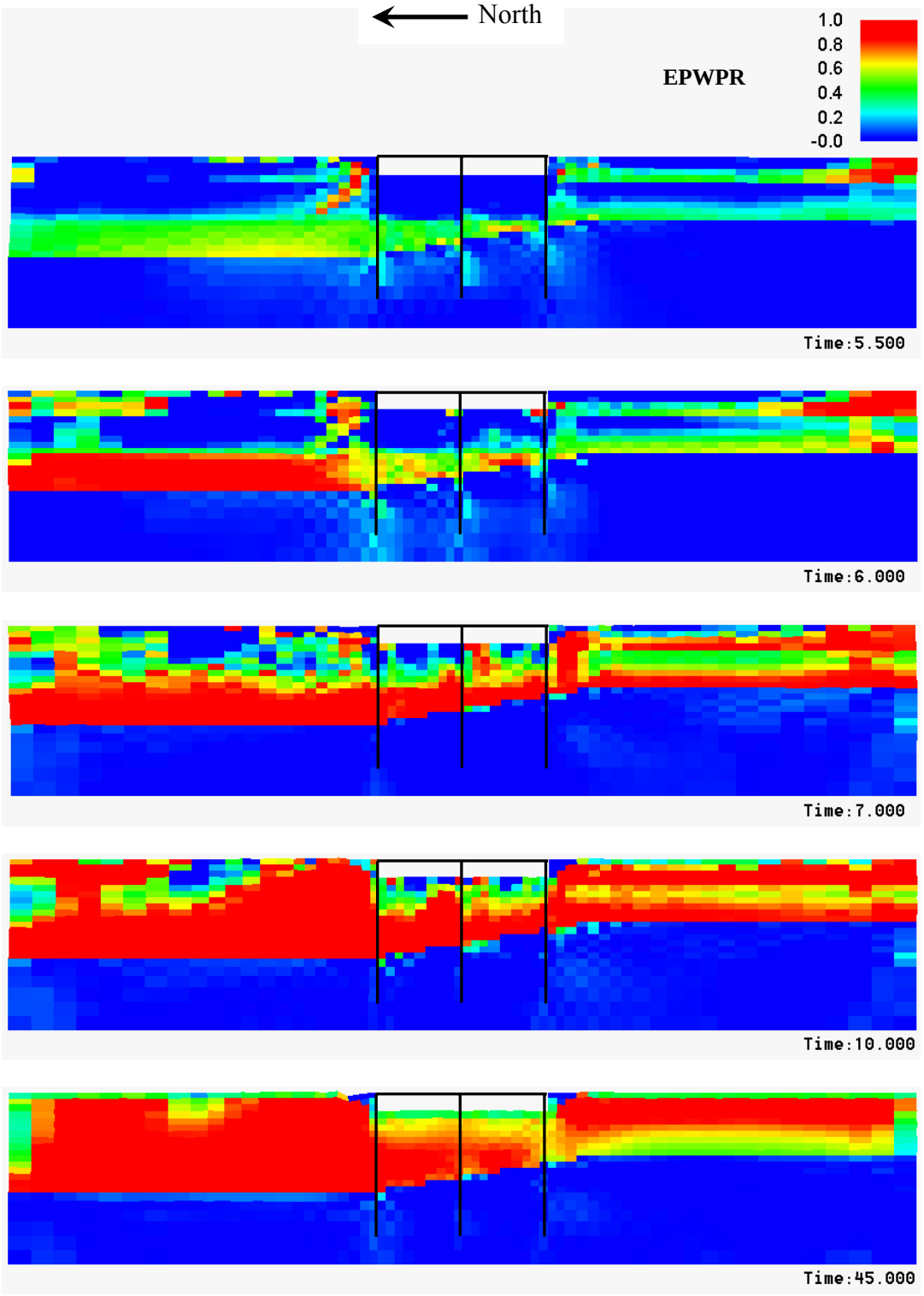


Figure 14-5: Development of excess pore water pressures and eventual liquefaction in the model during the deterministic analysis. Note that an excess pore water pressure ratio (EPWPR) of 1.0 indicates liquefaction and zero effective stress in the soil.

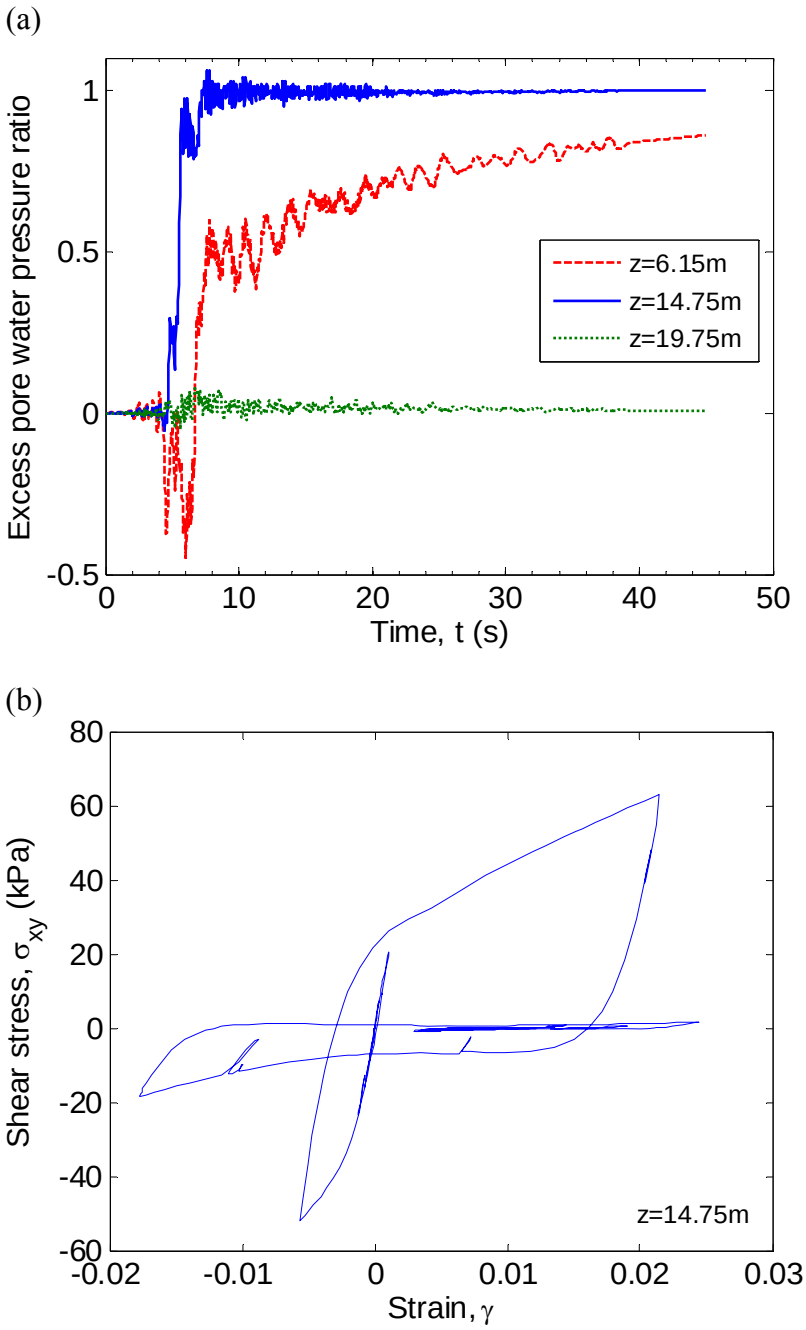


Figure 14-6: (a) Typical excess pore water pressure ratio development in the north free field (x=20 m in Figure 14-2) at various depths; and (b) typical shear stress-strain response.

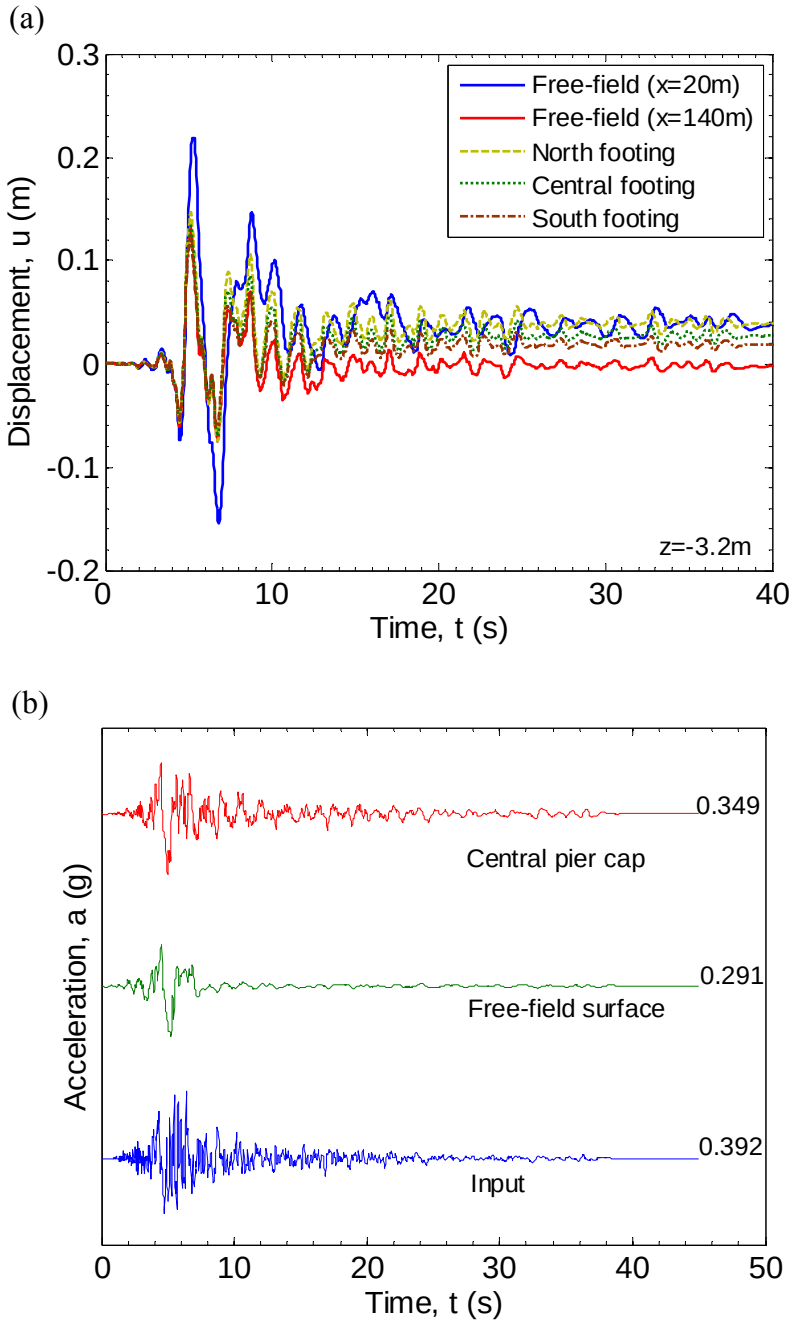


Figure 14-7: (a) displacement response history of the free field and at the pile footings; and (b) comparison of input, free-field surface, and pier cap acceleration histories (maximum values given at the end of each history).

Figure 14-8a illustrates the bending moment profiles in the piles and abutments/pier at $t = 5.15$ s which corresponds to the peak footing displacements in Figure 14-7a. It can be seen that the seismic demand on the pile foundations is significant with both north and central piles exceeding their respective yield moments, and the south piles exceeding the cracking moment. The variation in the $N_1 = 10$ to $N_1 = 30$ boundary depth (e.g. Figure 14-2) is also observed to have a pronounced effect on the depth at which the peak negative bending moment is

developed in the piles. The effect of this depth variation also causes larger soil displacements on the north side of the model relative to the south. As the large axial stiffness of the bridge superstructure effectively enforces equal displacements of the top of the abutments (with the exception of seating displacement discussed in the next paragraph), this variation in soil displacements in the horizontal direction also causes significantly different moments in the upper half of the piles and the abutments/pier. Figure 14-8b illustrates the shear force histories for the two abutments and central pier. It is immediately evident that forces in the north and south abutments are of opposite sign indicating that the bridge superstructure is predominately restraining the displacements of the north abutment/piles (where soil displacements are relatively large), and increasing the displacement of the south abutment/piles (where soil displacements are relatively smaller).

Figure 14-9 illustrates the relative displacement between the bridge superstructure and abutment (herein referred to as seating displacement) at the north and south abutments (the superstructure is fixed to the central pier). The maximum seating displacement of 10 cm at the south abutment is significantly large to require some post-earthquake inspection and repair (i.e. Table 14-4). In addition, explicitly accounting for the seating displacement restricts the maximum shear force which can be transmitted between the bridge superstructure and the north and south abutments, which was observed to reduce the bending moments in the abutments relative to those in the central pier.

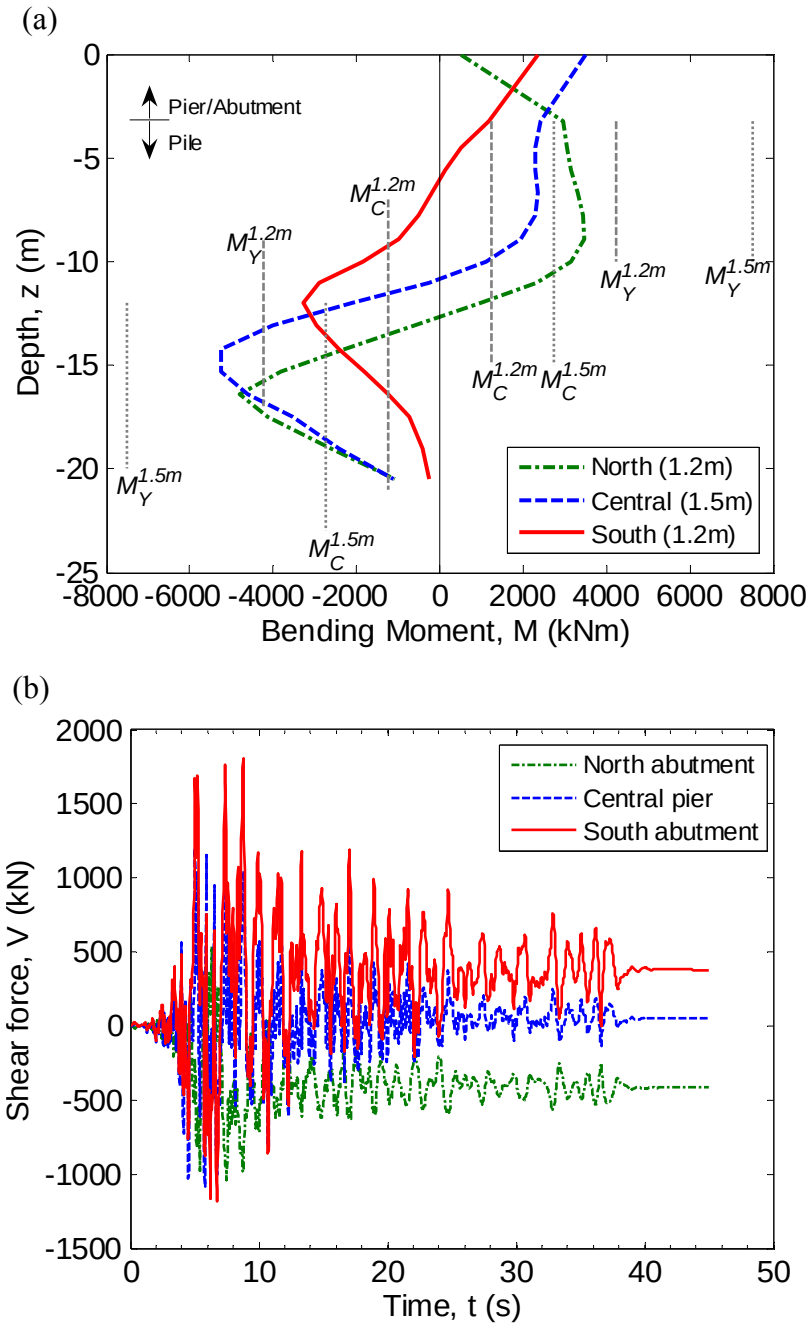


Figure 14-8: (a) bending moment profiles of the pile foundations at t=5.15s; and (b) shear force time histories in the abutments/pier. Symbols $M_C^{1.2m}$, $M_Y^{1.2m}$, $M_U^{1.2m}$, and $M_C^{1.5m}$, $M_Y^{1.5m}$, $M_U^{1.5m}$ represent cracking, yielding, and ultimate moment capacities for the 1.2m and 1.5m piles, respectively.

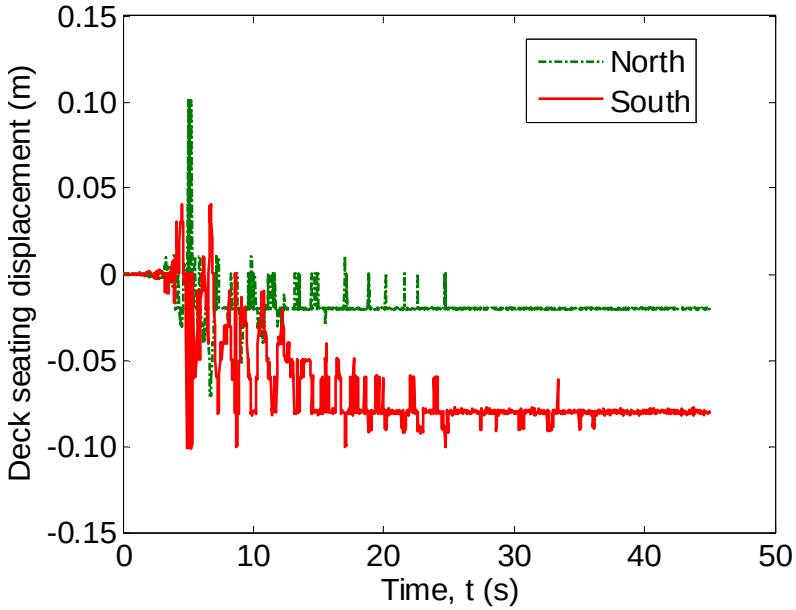


Figure 14-9: Deck seating displacement at the north and south abutments.

14.7 Probabilistic seismic response analyses

Clearly a vast amount of information and insight into the seismic response of the entire bridge-foundation-soil system is possible by rigorously examining such seismic effective stress analyses discussed in the previous section. However, the results of a single seismic effective stress analysis do not allow for explicit consideration of the rate of exceedance and uncertainties in the characteristics of the incident ground motion or the consequences associated with the seismic response. Consideration of the uncertainties associated with the incident ground motion can be accounted for by subjecting the developed numerical model of the system to multiple ground motions of various levels of intensity. This section discusses the results of these *probabilistic seismic response analyses* using the aforementioned ground motions selected from seismic hazard deaggregation.

Based on the observations of various deterministic analyses, a total of nine different engineering demand parameters (EDP's) were monitored in each of the analyses discussed in this section. These EDP's were: peak curvature throughout the length of each of the three piles; peak curvature in the abutments and central pier; maximum seating displacement at the two abutments; and the maximum value of the vertical settlement of the bridge approach embankments. As discussed in the previous sections, peak ground acceleration (PGA) was used as the ground motion intensity measure.

14.7.1 Piles and abutments/pier

Figure 14-10 illustrates the results of the seismic response analyses for twenty ground motions at nine intensity levels for peak curvature in the north, central, and south piles; north and south abutments; and central pier. Several points are worthy of note in Figure 14-10. Firstly, as expected, the mean seismic demand for each of the EDP's increases with an increase in the input ground motion intensity. Secondly, there is a large amount of variation in the magnitude of the results for a given value of *PGA* (e.g. for $PGA = 0.46$ g the peak curvature in the north pile ranges from 0.0004 to 0.005). This large variation occurs because of the aforementioned inefficiency of *PGA* as a ground motion intensity measure for the seismic response of soft soil deposits (but the fact that seismic hazard for superior intensity measures do not exist for this site). Thirdly, there is a significant variation in the level of seismic demands for the different pile and abutment/pier components depending on their location in the bridge-foundation-soil system.

As observed in the deterministic analysis discussed in the previous section, the curvature demand for the north piles is larger than the other two piles because of the larger soil displacements at this location due to the variation in the depth of the $N_1 = 10$ to $N_1 = 30$ soil layer boundary. The central pier is also observed to have a higher curvature demand than that in the north and south abutments. This higher demand in the central pier occurs primarily due to the fixed connection between it and the superstructure, while sliding of the abutment-superstructure connections limits the maximum shear forces, and hence moment, in the abutments.

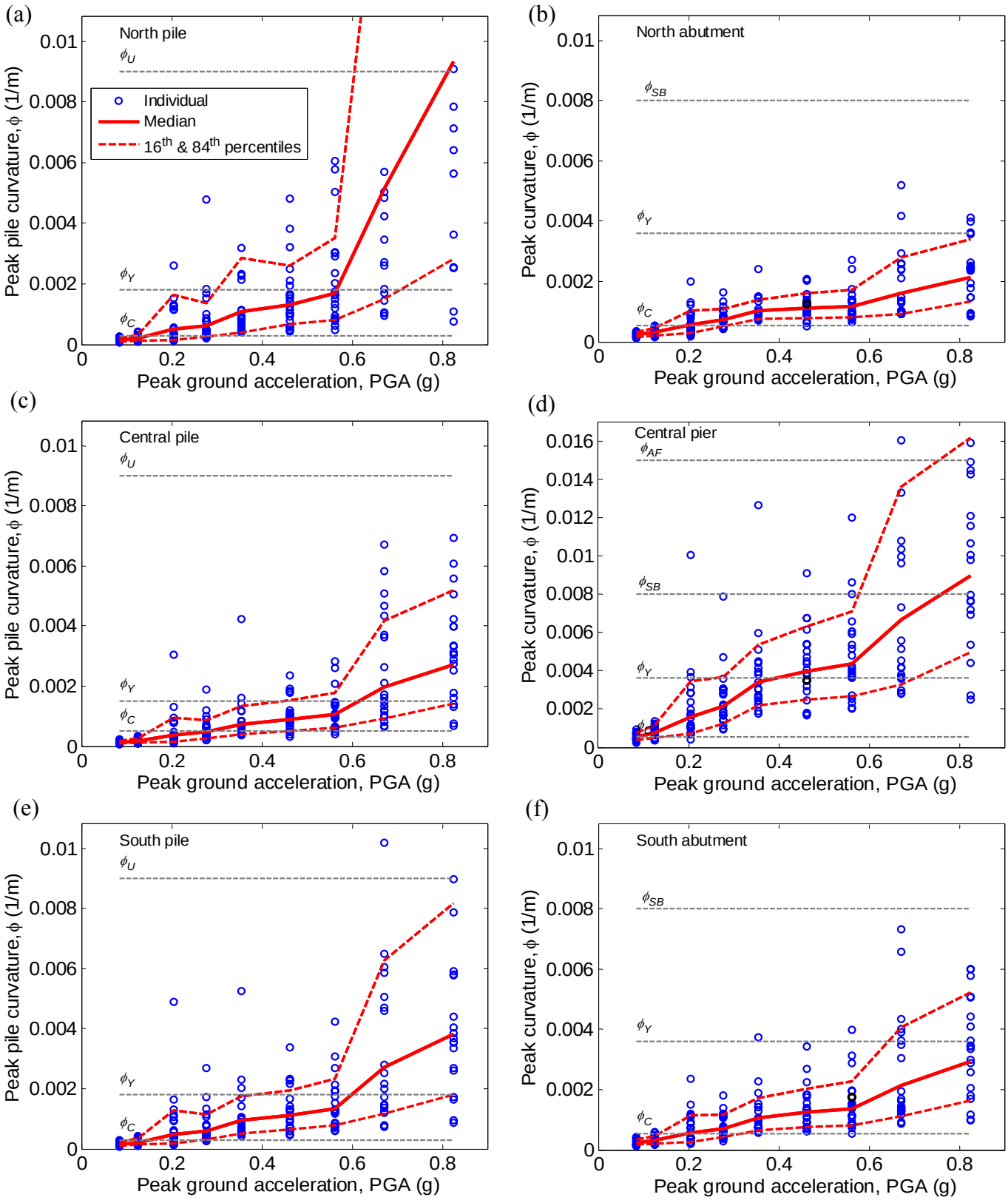


Figure 14-10: Probabilistic seismic response analysis results for the pile foundations, and abutments/pier at the north, central and south locations. Median values of various damage states annotated are given in Table 14-4.

14.7.2 Deck seating displacement

Figure 14-11 illustrates the results of the probabilistic seismic response analyses for the maximum seating displacements at the north and south abutments. As noted in regard to Figure 14-10 it is again observed that there is significant scatter in the peak seating displacements for a given *PGA*. It can be seen that even for small levels of ground motion (PGA > 0.2g) there is a significant likelihood of the maximum seating displacement exceeding the median value for minor repair (i.e. Table 14-4). The likelihood of complete unseating failure is however significantly low over all levels of PGA considered.

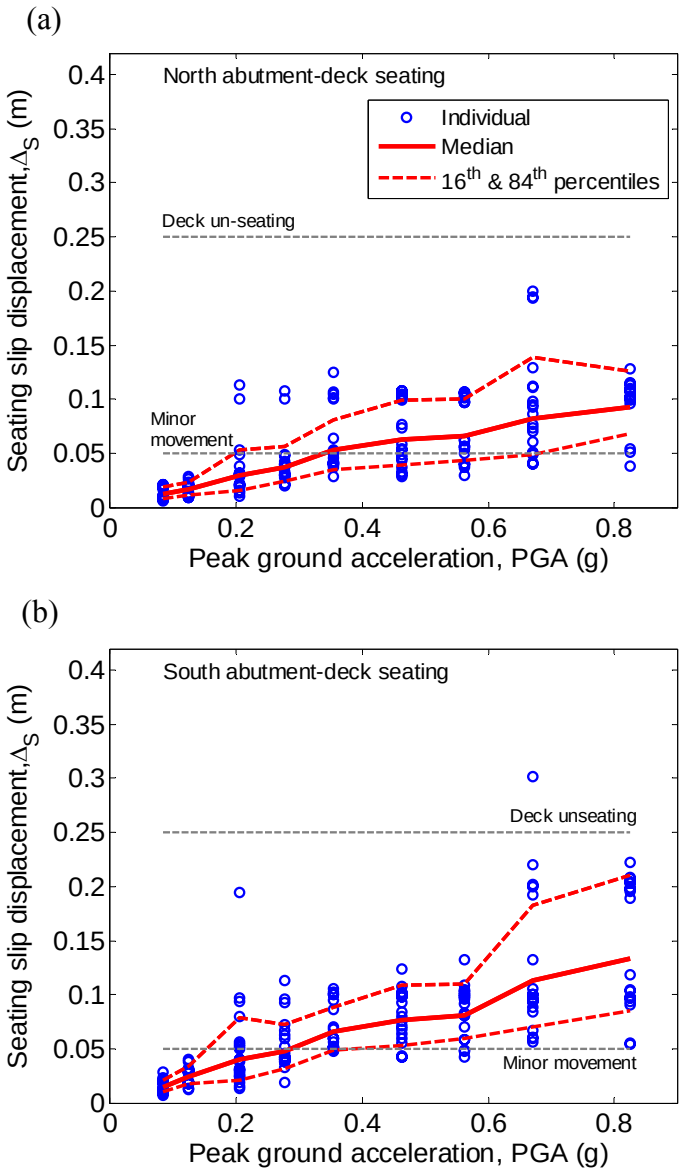


Figure 14-11: Probabilistic seismic response analysis results for the relative displacement between abutments and bridge deck. Median values of various damage states annotated are given in Table 14-4.

14.7.3 Approach settlements

Significant liquefaction of foundation soils and lateral spreading can cause large settlements of the ground near the approaches to a bridge superstructure. Figure 14-12 illustrates the maximum vertical settlements computed at the approaches to either side of the bridge superstructure. It can be seen that significant settlements of up to 0.4 m were computed, and also that the dispersion in the magnitude of the settlement is significant. Ground motions above 0.2g PGA have a high likelihood of causing vertical settlements which will require minor post earthquake repair, while significant approach settlements requiring major earthworks are likely only under large ground motions (i.e. $PGA > 0.6$g), with the exception of some 'outlying' responses at 0.355g PGA.

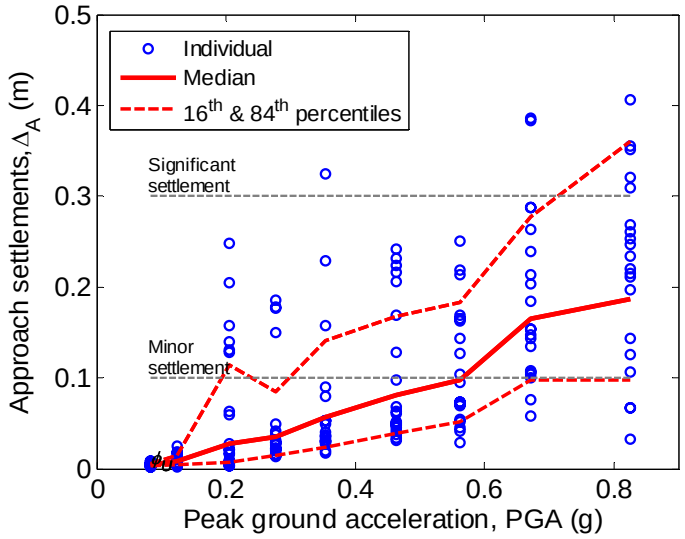


Figure 14-12: Probabilistic seismic response analysis results for the maximum vertical settlement at the bridge approaches. Median values of various damage states annotated are given in Table 14-4.

14.8 Seismic demand hazard

By combining the seismic response analyses obtained in the previous section, which account for the variability in response due to complex ground motion excitation, with the seismic hazard curve, describing the likelihood of various levels of ground motion, it is possible to compute the annual rate of exceeding various levels of demand for each of the different EDP's monitored. Mathematical and computational details of computing the annual rate of exceeding various levels of demand can be found in, for example, Bradley *et al*. [2].

Figure 14-13a illustrates the annual rate of exceeding various levels of peak pile

curvature for each of the three piles in the computational model. The effect of the variation in demand for the piles observed in Figure 14-10 is also apparent in Figure 14-13a. Based on the monotonic moment-curvature relationship of the piles the median values of the cracking, yielding, and ultimate damage states (Table 14-4) are also given in Figure 14-13a. It can be seen that the north and south piles are more vulnerable (i.e. have higher damage state exceedance rates) than the larger 1.5m diameter central pile, with the north pile particularly vulnerable for higher levels of curvature. Figure 14-13b illustrates the annual rate of exceeding various levels of peak curvature of the abutments and central pier. As observed in Figure 14-10, it can be seen that the demand on the central pier is significantly greater than that for the north and south abutments, with the central pier having annual damage state exceedance frequencies typically an order of magnitude larger than the abutments.

Figure 14-14a illustrates the annual rate of exceeding various levels of peak seating displacement at the north and south abutments. It can be seen that the annual rate of exceedance for various levels of demand at both the north and south abutments is quite similar. The exceedance rates for the displacement corresponding to the median value for minor repair (Table 14-4) are slightly above $1x10^{-3}$, while those for deck unseating are approximately $1x10^{-5}$. Figure 14-14b illustrates the annual rate of exceeding specified levels of approach embankment vertical settlement. It can be seen that the exceedance rates for the median values corresponding to minor and major repair are quite large (approximately $1x10^{-3}$ and $1x10^{-4}$, respectively).

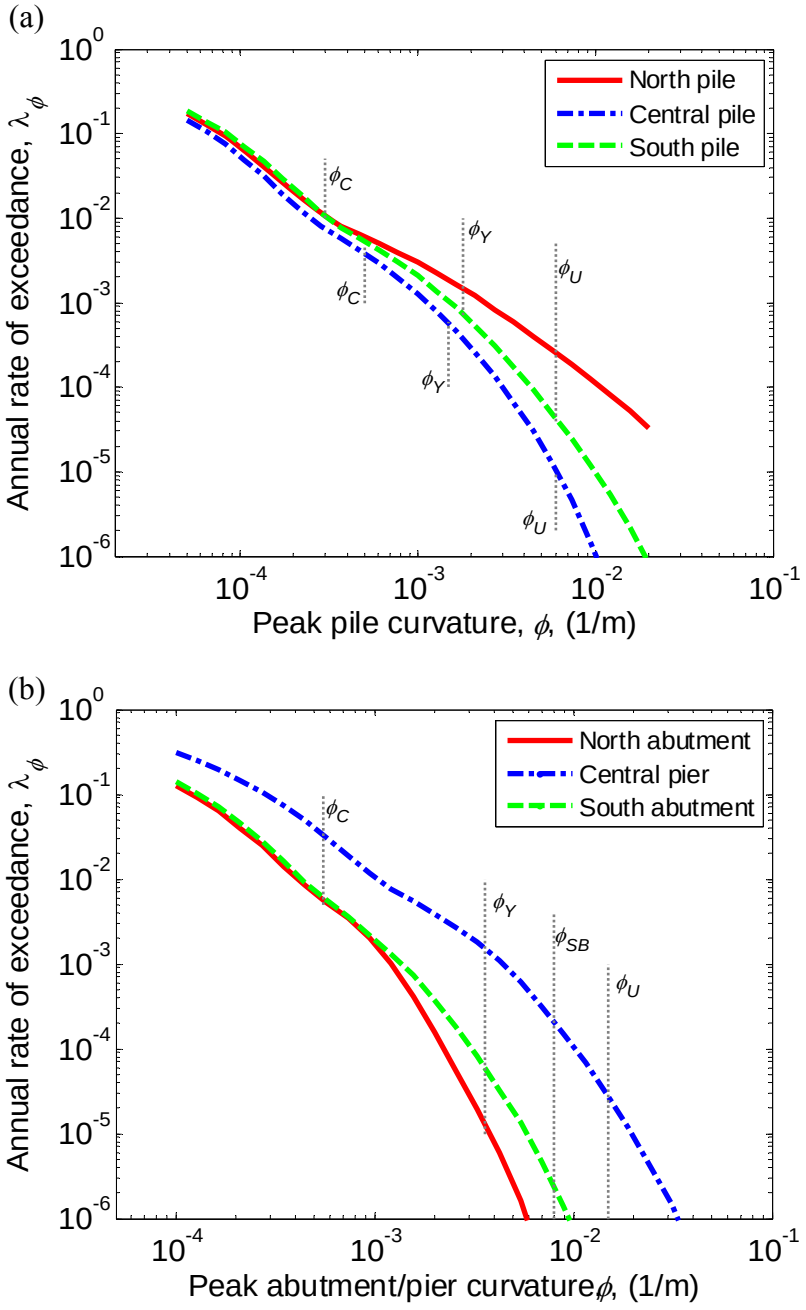


Figure 14-13: demand hazard curves for: (a) peak pile curvature; and (b) peak abutment/pier curvature. Median values of various damage states annotated are given in Table 14-4.

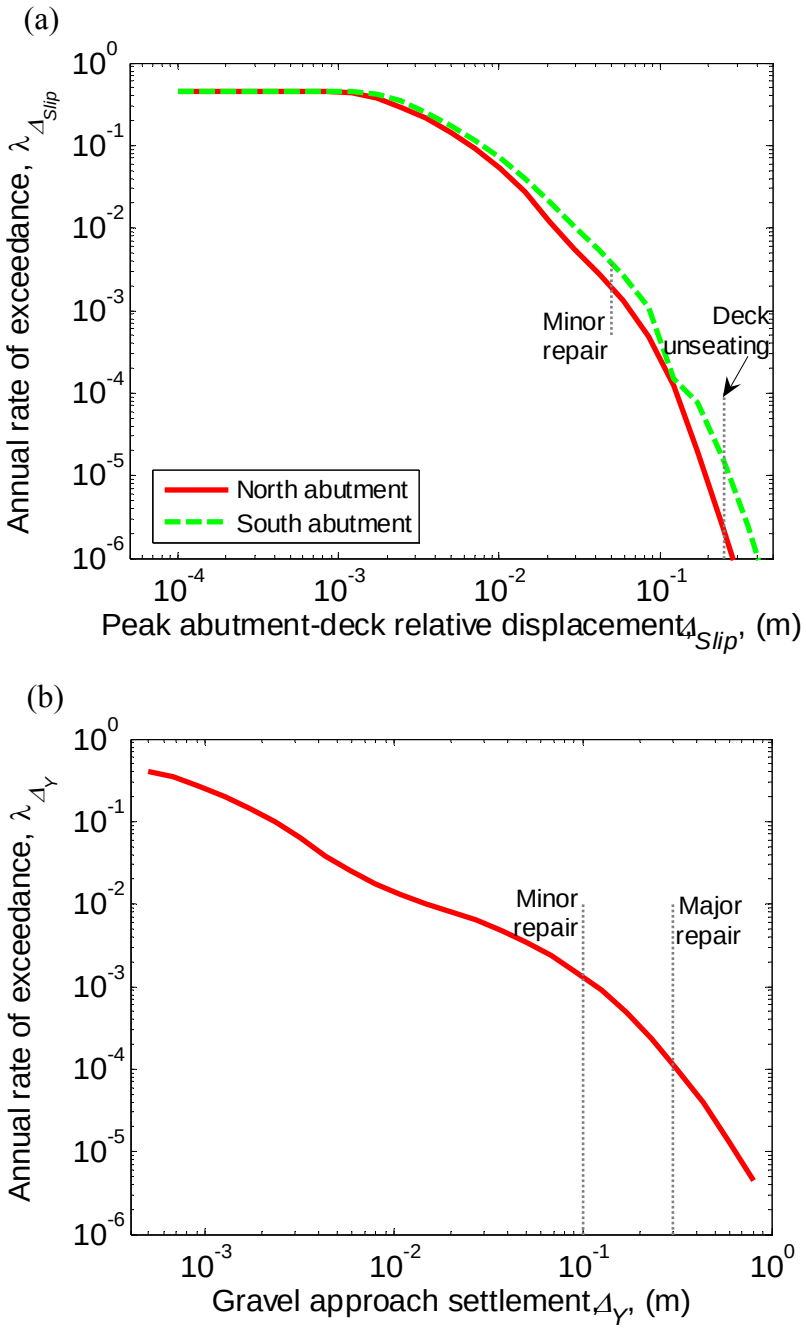


Figure 14-14: demand hazard curves for: (a) peak seating displacement; and (b) peak approach settlement. Median values of various damage states annotated are given in Table 14-4.

14.9 Probabilistic seismic loss assessment

The seismic risk of the bridge-foundation-soil system in terms of the explicit loss consequences due to structural response may be viewed as ultimate measures of seismic performance for decision making. In order to conduct such a seismic loss assessment, the consequences, in the form of direct repair cost and repair duration, due to various states of

damage for each of the components of the system are required. To this end, a professional cost estimator was engaged to develop cost estimates and repair durations due to various levels of damage in each of the components of the Fitzgerald bridge [25], a summary of which is given in Table 14-4. Particular points of note in Table 14-4 are: (i) no repair actions are performed for cracking in the pile foundations; (ii) the ultimate curvature capacity of the piles is less than that given by a monotonic moment-curvature analysis (i.e. Table 14-1) because of cyclic degradation effects; (iii) because the repair of cracking in the central pier can be done without removing the abutment backfills it causes less delay than repairing the abutments.

A recent network traffic flow analysis [26] which estimated the travel delay and vehicle running costs due to inoperability of the Fitzgerald Avenue twin bridges at $10,720/day was used to convert repair durations to loss of functionality costs.

14.9.1 Loss assessment framework

The loss assessment presented herein employed the seismic loss assessment tool, SLAT and is based on the PEER framing formula [4]. The treatment of the direct repair losses is the same as that of Bradley *et al*. [2] and Aslani [1], and is only briefly mentioned here. The treatment of repair duration losses presented here is novel and is discussed in greater detail below.

Each of the components of the bridge-foundation-soil system (i.e. piles, abutments) are denoted as performance groups (PG's), and for a given ground motion intensity the expectation and variance in the loss to each performance group is defined as [e.g. 1, 2]:

$$\mu_{L_i|IM}(im) = \int \mu_{L_i|EDP}(edp) f_{EDP|IM}(edp|im) dEDP \tag{14-1}$$

$$\sigma^2_{L_i|IM}(im) = \int \left[\mu^2_{L_i|EDP}(edp) + \sigma^2_{L_i|EDP}(edp)\right] f_{EDP|IM}(edp|im) dEDP - \mu^2_{L_i|IM}(im) \tag{14-2}$$

where $\mu_{L_i|IM}(im)$ and $\sigma^2_{L_i|IM}(im)$ are the expected value and variance in the loss (either direct repair cost or repair duration) of PG i given $IM = im$; $\mu_{L_i|EDP}(edp)$ and $\sigma^2_{L_i|EDP}(edp)$ are the mean and variance in the loss of PG i given $EDP = edp$, respectively; and $f_{EDP|IM}(edp|im)$ is the probability density function (pdf) for the *EDP|IM* relationship. The mean and variance in the loss of PG i given $EDP = edp$ are obtained by combining the damage fragility and loss data given in Table 14-4 using the total probability theorem [e.g. 1,

2], while $f_{EDP|IM}\left(edp|im\right)$ is obtained from the results of the probabilistic seismic response analyses (e.g. Figure 14-10-Figure 14-12).

Table 14-4: Damage states of components with fragility and loss data.

Component	Damage State	Fragility $\mu_{EDP\|DS}$, $\sigma_{\ln EDP\|DS}$	Repair	Repair cost ($M) $\mu_{L_{RC}\|DS}$, $\sigma_{\ln L_{RC}\|DS}$	Repair duration (days) $\mu_{L_{DT}\|DS}$, $\sigma_{\ln L_{DT}\|DS}$
1.2m diameter pile	Cracking	0.0003/m, 0.15	None	0.0 , 0.0	0.0 , 0.0
	Yielding	0.0018/m, 0.25	Replace pile with fixed connection to pile cap	0.18 , 0.40	120.0 , 0.40
	Failure	0.0060/m, 0.50	Loss of vertical carrying capacity and significant settlement. Replace pile and repair settled abutment	0.48 , 0.40	180.0 , 0.40
1.5m diameter pile	Cracking	0.0005/m, 0.20	None	0.0 , 0.0	0.0 , 0.0
	Yielding	0.0015/m, 0.25	Replace pile with fixed connection to pile cap	0.20 , 0.40	120.0 , 0.40
	Failure	0.0060/m, 0.50	Loss of vertical carrying capacity and significant settlement. Replace pile and repair settled abutment	0.50 , 0.40	180.0 , 0.40
Abutments	Cracking	0.00055/m, 0.15	Epoxy injection. Requires excavation of approach to access	0.02 , 0.35	90.0 , 0.35
	Yielding	0.0036/m, 0.25	Externally reinforce	0.18 , 0.35	180.0 , 0.35
	Spalling/Buckling	0.0080/m, 0.35	Replace abutment	0.50 , 0.35	240.0 , 0.35
	Axial Failure	0.0150/m, 0.50	Replace abutment and deck span	1.00 , 0.35	360.0 , 0.35
Central pier	Cracking	0.00055/m, 0.15	Epoxy injection. Requires excavation of approach to access	0.02 , 0.35	20.0 , 0.35
	Yielding	0.0036/m, 0.25	Externally reinforce	0.18 , 0.35	180.0 , 0.35
	Spalling/Buckling	0.0080/m, 0.35	Replace abutment	0.50 , 0.35	240.0 , 0.35
	Axial Failure	0.0150/m, 0.50	Replace abutment and deck span	1.00 , 0.35	360.0 , 0.35
Deck	Minor movement	0.05m, 0.20	Minor repair of expansion joints, mechanical devices for additional seating length	0.15 , 0.40	90.0 , 0.40
	Unseating	0.25m, 0.20	Deck unseats and falls, replace damaged deck	0.40 , 0.40	180.0 , 0.40
Approach embankments	Minor settlement	0.10m, 0.30	Cracking of approach road, requires new subgrade and basecourse	0.15 , 0.40	120.0 , 0.40
	Significant settlement	0.30m, 0.30	Earthworks to increase approach height and densify soil, new approach roading	0.60 , 0.40	180.0 , 0.40

The total direct repair loss of the bridge-foundation-soil system is assumed [1, 3, 5] to be the simple summation of the losses due to each PG. Thus the mean and variance in the total direct repair cost can be given by [e.g. 1, 2]:

$$\mu_{L_R|IM,NC}\left(im\right)=\sum_{i=1}^{N_{PG}}\mu_{L_i|IM}\left(im\right) \tag{14-3}$$

$$\sigma^2_{L_R|IM,NC}(im) = \sum_{i=1}^{N_{PG}} \sigma^2_{L_i|IM}(im) + 2\sum_{i=1}^{N_{PG}} \sum_{j=1}^{i-1} \rho_{L_i,L_j|IM}(im)\sigma_{L_i|IM}(im)\sigma_{L_j|IM}(im) \quad (14\text{-}4)$$

where N_{PG} is the number of performance groups and $\rho_{L_i,L_j|IM}(im)$ is the correlation between the loss given $IM = im$ for PG's i and j [27].

Unlike direct repair costs, it is unreasonable to simply assume that the total repair time to fix the entire bridge-foundation-soil system will be the sum of the repair durations for each component, since components can be fixed in parallel. To facilitate the computation of the repair duration ('downtime'), the idea of repair groups (RG's) is introduced here. A repair group has two distinct parts: (i) various PG's comprising the RG which must be repaired in a serial manner, and (ii) various other precursor RG's which must be first completed, before work on the particular repair group under consideration may commence. Mathematically speaking the time to complete repairs in repair group i given IM, $L_{RG_i}|IM$, is given by:

$$L_{RG_i}|IM = \sum_{j=1}^{N_{PG,RG_i}} L_j|IM + \max_{j=1:N_{RG,pre}} \left[L_{RG_j}|IM\right] \quad (14\text{-}5)$$

where N_{PG,RG_i} is the number of PG's in RG_i; and $N_{RG_i,pre}$ is the number of precursory RG's for RG i. Defining RG's as being comprised of various PG's as well as various precursor RG's has intentionally been made similar to typical Gantt chart construction scheduling software, such as Microsoft Project [28].

For the Fitzgerald Avenue twin bridges a total of five repair groups were considered as indicated in Table 14-5. RG_1 to RG_3 correspond to repair of components at the North, central, and South regions of the bridge; RG_4 to the approach embankments; and RG_5 to total repair of the bridge structure. The relatively simple configuration of the Fitzgerald Avenue twin bridges means the composition of the repair groups is not overly complicated, although the formulation of Equation (14-5) can easily handle more complex cases. It should be noted that because of the nature of Equation (14-5), it is not possible in general to estimate the statistical moments (i.e. mean and variance) of the repair duration for each of the repair groups in closed form, as was the case for direct repair cost (i.e. Equations (14-3) and (14-4)). Thus, after obtaining the statistical moments for all of the PG's using Equations (14-1) and (14-2), Monte Carlo simulation (with appropriate consideration of the correlation in the loss between multiple PG's [27]) was used to estimate the distribution of $L_{RG_i}|IM$. In particular, it should be noted that the expectation of the maximum of multiple random variables is not equal to the

maximum of their expectations (i.e. $E[\max(X_1,..,X_n)] \neq \max(E[X_1],..,E[X_n])$), as used in Mitrani-Reiser [29].

Table 14-5: Repair groups used in the repair duration computations.

RG number	Description	*PG's*	Precurssor RG's
1	North end of bridge	North piles, north abutment, north deck seating	None
2	Central portion of bridge	Central piles, central pier	None
3	South end of bridge	South piles, south abutment, south deck seating	None
4	Approach embankments	North/south approach embankments in parallel	None
5	Total repair time	None	RG_1, RG_2, RG_3, RG_4

14.9.2 *L|IM* results

Figure 14-15 illustrates the distribution of the direct repair loss and downtime for the bridge-foundation-soil system as a function of peak ground acceleration, *PGA*. It is observed that above 0.1g *PGA*, both measures of loss rise steeply with increasing *PGA*, and that there is significant uncertainty in the losses as illustrated by the difference between the mean ± one standard deviation values. It should be noted that as seismic response analyses were performed only up to 0.825g *PGA* (there was no interest in the seismic response of the system for rarer ground motions). This upper bound *PGA* value (which has an annual rate of exceedance of $\lambda_{IM} = 4.0 \times 10^{-5}$) should be borne in mind when examining the loss hazard results to follow.

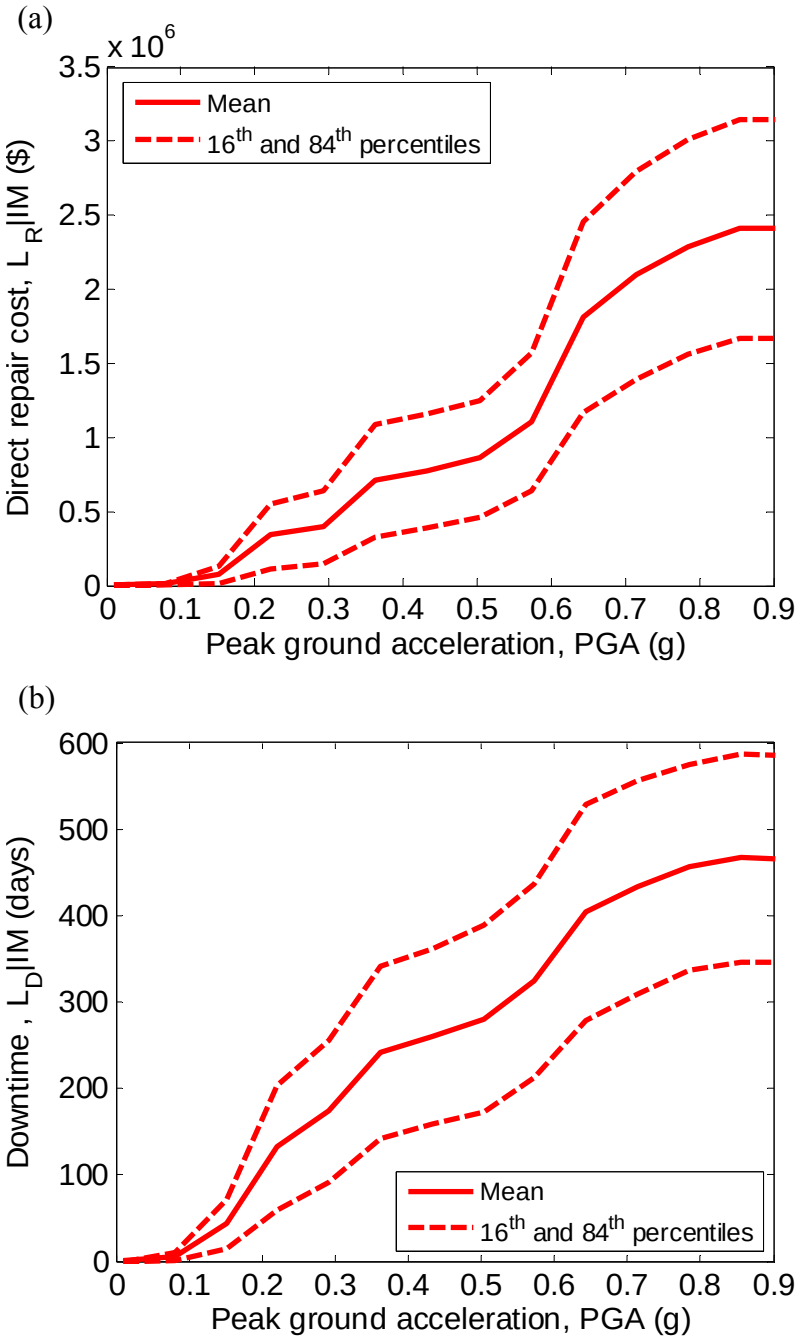


Figure 14-15: Distribution of: (a) direct repair cost; and (b) downtime, as a function of peak ground acceleration.

Figure 14-16 illustrates the deaggregation of the expected direct repair loss as a function of *PGA* for the nine different components which comprise the bridge-foundation-soil system. It can be seen that over a wide range of *PGA*, the direct repair loss is primarily attributed to damage to the north piles, central pier and liquefaction of the approach embankments. Conversely, the cost to repair damage to the north and south abutments comprise a significantly smaller proportion of the total repair costs.

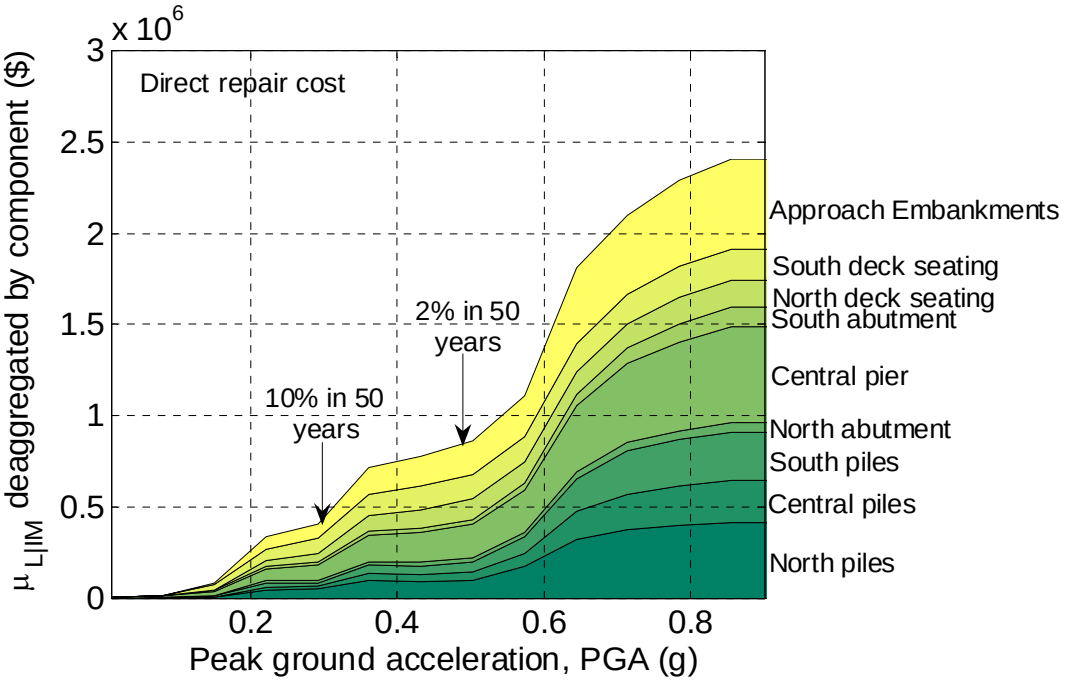


Figure 14-16: Deaggregation of the expected direct repair cost due to the different components of the bridge-foundation-soil system.

Figure 14-17a and Figure 14-17b illustrate the deaggregation of the expected direct repair cost for PGA intensities with 10% and 2% probabilities of exceedance in 50 years, respectively. The same features noted in regard to Figure 14-16 are also observed in Figure 14-17a and Figure 14-17b. In particular, unlike some previously conducted loss estimation case-studies [e.g. 2], Figure 14-17a and Figure 14-17b illustrate that there is a surprisingly small variation in the contribution of the different components at these two different levels of ground motion.

Figure 14-18a compares the expected downtime as a function of PGA for the four different repair groups and the total bridge downtime. It can be seen that RG_4 (approach embankments) are the repair group which is completed first over the entire range of ground motion intensity, with RG_1 and RG_3 taking the longest to complete over the majority of *PGA* values of interest. Figure 14-18b illustrates the dispersion (i.e. lognormal standard deviation) in the repair group downtimes as a function of *PGA*. As has been noted by others in regard to direct repair cost [30, 31], the dispersion in the repair group downtime decreases with increasing expected downtime values. Also, as mentioned in the previous section, it is noted that the expected total downtime is significantly larger than the expected downtime of the four repair groups, while the dispersion in the total downtime is less than that of the individual repair groups.

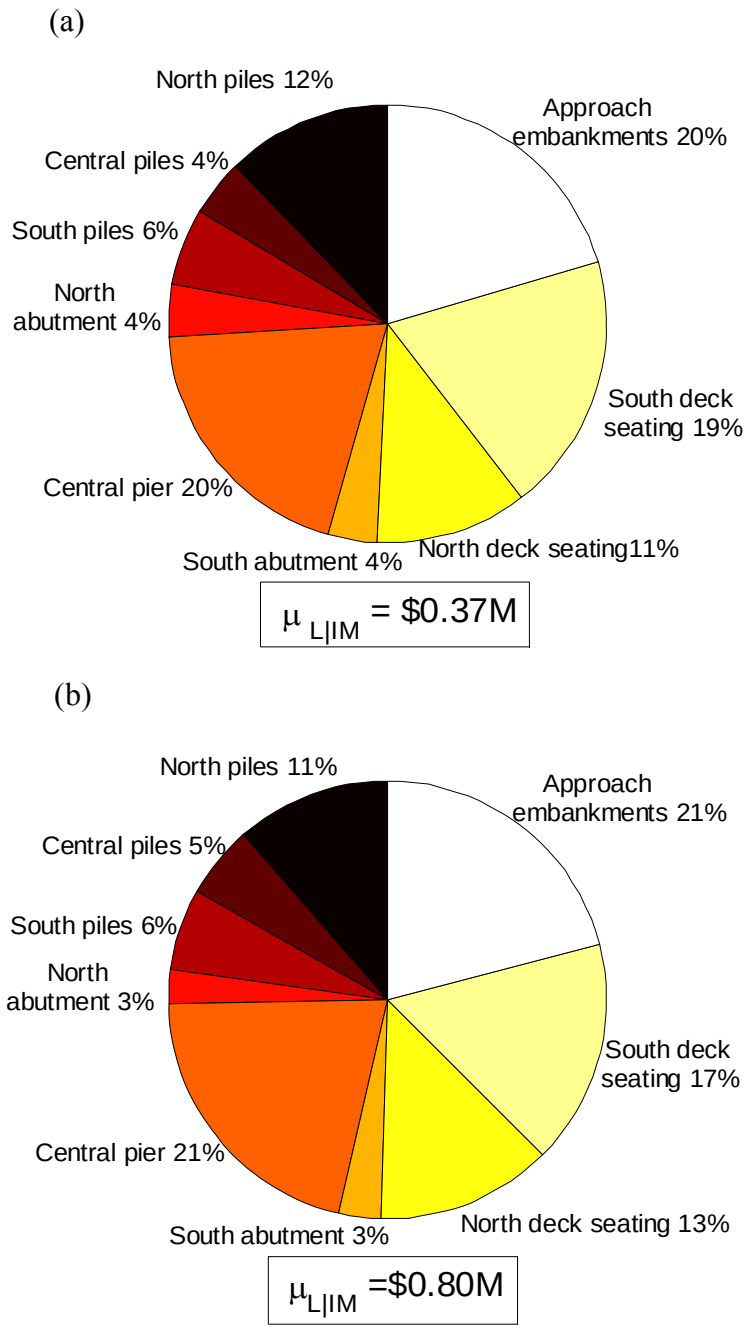


Figure 14-17: Deaggregation of the expected direct repair cost due to the different components of the bridge-foundation-soil system for (a) PGA = 0.27g (10% in 50 years) and (b) PGA = 0.46g (2% in 50 years).

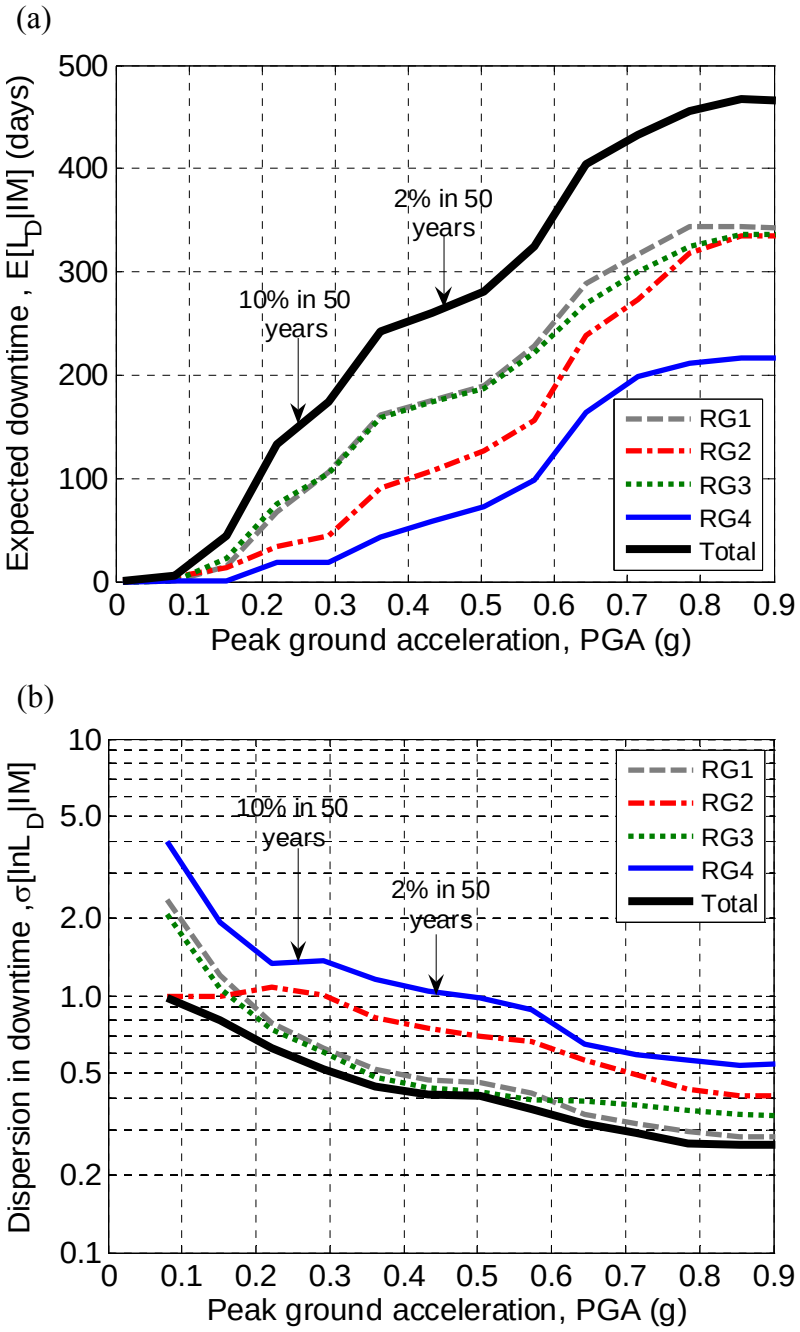


Figure 14-18: (a) expected value; and (b) dispersion, in the downtime for the various repair groups of the bridge-foundation-soil system as a function of ground motion intensity.

Figure 14-19a and Figure 14-19b illustrate the deaggregation of the repair group downtime at the PGA intensities with 10% and 2% probability of exceedance in 50 years, respectively. It can be seen that as noted in regard to Figure 14-18, repair of the approach embankments is not critical for either of these PGA intensities. It is worthy of note that Figure 14-19a illustrates the total repair time for RG_2 (in particular, the central pier), is small considering the significance of the central pier in the deaggregation of the direct repair costs

(e.g. Figure 14-17). This is because, as noted in Table 14-4, the duration required to repair cracking (using epoxy injection) in the central pier does not require excavation of the gravel backfills as in the case of repairing cracking in the bridge abutments. As for the direct repair cost, the downtime to repair damage in the north piles is larger than that for the central and south piles. The time to re-establish adequate seating length of the bridge deck at both the north and south abutments is also an important contributor to the total expected downtime of the bridge-foundation-soil system.

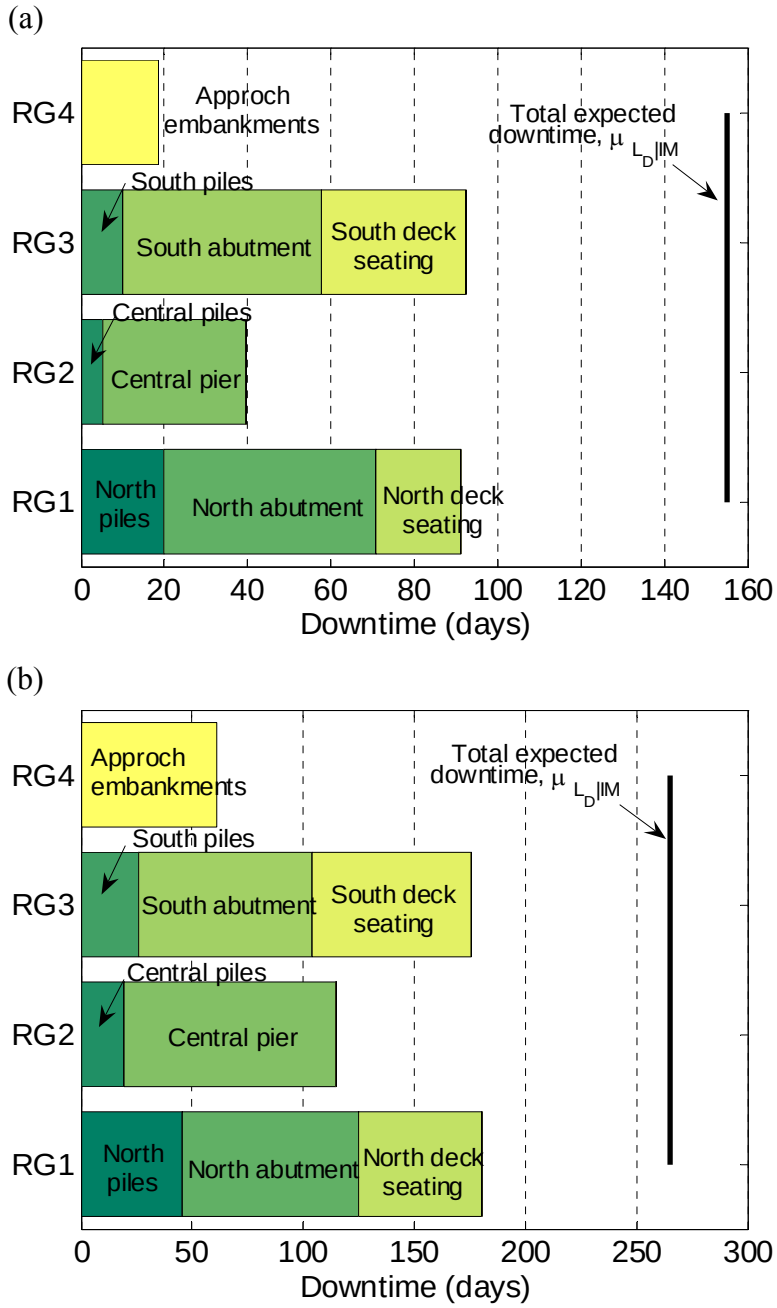


Figure 14-19: Deaggregation of the expected downtime due to the different components of the bridge-foundation-soil system for (a) PGA = 0.27g (10% in 50 years) and (b) PGA = 0.46g (2% in 50 years).

14.9.3 Loss hazard results

By combining the distribution of loss given ground motion intensity with the ground motion hazard, the annual rate of exceeding some level of loss can be computed as:

$$\lambda_L(l) = \int G_{L|IM}(l|im) \left| \frac{d\lambda_{IM}(im)}{dIM} \right| dIM \tag{14-6}$$

where $\lambda_L(l)$ is the annual rate of the loss, L, exceeding l, $G_{L|IM}(l|im)$ is the probability of $L > l$ given $IM = im$, and $\lambda_{IM}(im)$ is the ground motion hazard (i.e. Figure 14-3a). $G_{L|IM}(l|im)$ is obtained based on the statistical moments of the $L|IM$ distributions obtained in the previous section and using the common assumption of lognormality for $L|IM$ [27, 30] (which was verified for this particular case-study, but is omitted for brevity).

Figure 14-20 illustrates the annual rate of exceeding various levels of downtime for the four different repair groups and for the total system. While the cracking damage state in the central pier is the primary contributor to the total downtime for values of less than 10 days, for more significant events, the total downtime is predominantly due to RG_1 and RG_3.

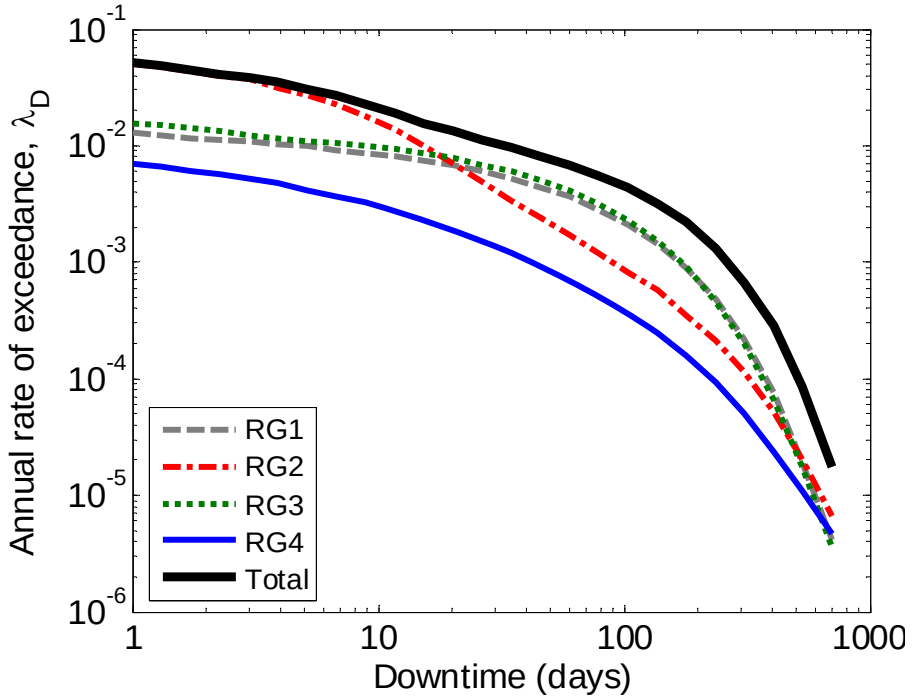


Figure 14-20: Downtime risk curve for the four repair groups and the entire bridge-foundation-soil system.

Knowing that travel delay and vehicle running costs due to inoperability of the bridge structure amount to $10,720/day [26] (assumed to be a deterministic value), the annual rate of

exceeding a specified level of downtime (in days) can be converted to the annual rate of exceeding a specified level of economic loss. Figure 14-21 illustrates the annual rate of exceeding some level of economic loss due to inoperability of the bridge, as well as the annual rate of exceeding some level of direct repair cost of the system. It can be seen that over the full range of economic losses (or annual rates of exceedance), the economic implications due to loss of functionality is significantly larger than that due to direct repair of damage. In particular, the 10% in 50 year exceedance probability ($\lambda_L = 2.1\text{x}10^{-3}$) direct repair and loss of functionality losses are \$0.475M and \$2.03M, while the 2% in 50 year exceedance probability ($\lambda_L = 4.0\text{x}10^{-4}$) losses are \$1.20M and \$3.95M respectively. For comparative purposes it is again noted that the book-value of the Fitzgerald Avenue twin bridges is only \$2.4M. Thus, there is a 2% in 50 year probability that the total loss will exceed almost \$5.2M, over two-times the book-value of the infrastructure itself.

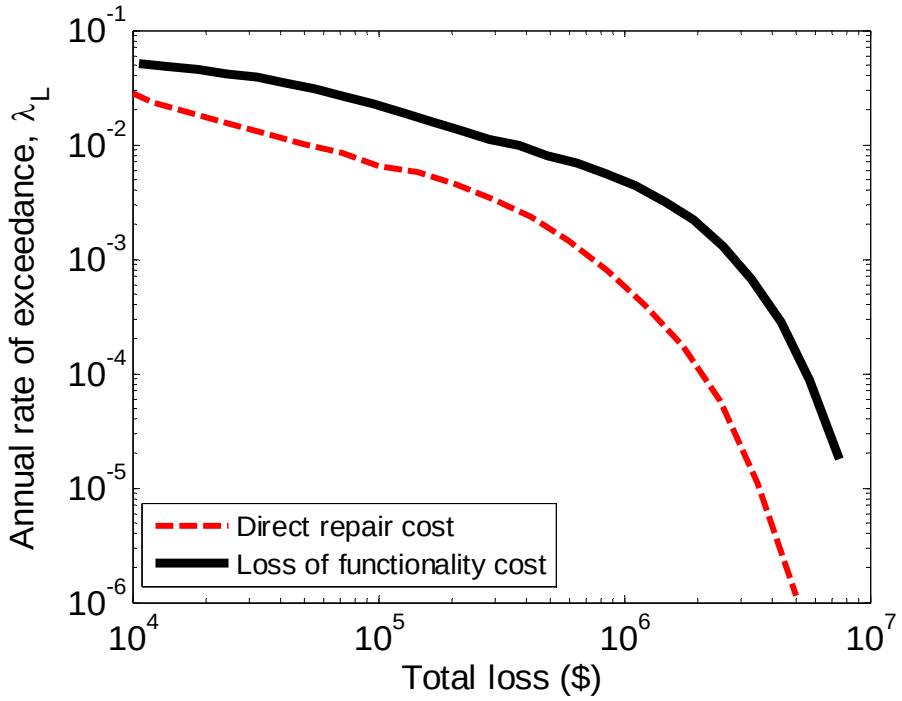


Figure 14-21: Comparison of the annual rate of exceedance of losses due to direct repair cost and loss of functionality.

14.10 Conclusions

This paper has presented a holistic seismic performance and loss assessment of the Fitzgerald Avenue Twin bridges, located on a major arterial route of Christchurch, New Zealand. A finite element model of the longitudinal direction of the bridge-foundation-soil system was constructed, and detailed deterministic seismic effective stress analyses were used

to gain considerable insight into the complex seismic response of the system involving significant nonlinear dynamic soil behaviour and soil-structure-interaction. Based on the information that they provide such detailed deterministic analyses should always accompany a subsequent probabilistic assessment of seismic performance.

To explicitly account for the uncertainty in the intensity and other characteristics of earthquake-induced ground motions, 20 different input ground motion records were selected based on seismic hazard deaggregation for each of the nine different PGA levels considered. The statistics of the seismic response due to multiple ground motions at various levels of *PGA* were combined with the *PGA* ground motion hazard to obtain the annual rate of exceeding various levels of seismic demand for various components of the bridge-foundation-soil system.

Fragility functions for various damage states of the system components were developed, and professional cost estimates were obtained in terms of direct repair cost and repair duration for each of these damage states. A detailed probabilistic seismic loss assessment of the bridge-foundation-soil system was performed. Through the use of loss deaggregation it was determined that the north piles, central pier, deck seating, and approach embankments are the most vulnerable components of the system. The consequences due to loss of functionality of the bridge during repair were significantly larger than the direct repair costs. The fact that there is a 2% in 50 year probability of combined direct repair and loss of functionality costs exceeding \$5.2M, more than two times the \$2.4M book-value of the asset, indicates the significant level of risk that seismic hazards pose to this infrastructure.

14.11 Acknowledgements

Financial support from the New Zealand Tertiary Education Commission and the New Zealand Earthquake Commission as well as assistance in obtaining cost estimation data from Mr. Lloyd Greenfield (Christchurch City Council) and Mr. William Hopkins (Hopkins Engineering) are greatly appreciated.

14.12 References

[1] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.

[2] Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, and Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society for Earthquake Engineering* 2009; **42**(2): 96-110.

[3] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1973-1997, DOI: 10.1002/eqe.694.

[4] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).

[5] Porter KA, Kiremidjian AS, and LeGrue JS. Assembly-Based Vulnerability for Buildings and Its Use in Performance Evaluation. *Earthquake Spectra* 2001; **17**(2): 291-312.

[6] Tonkin and Taylor Ltd. Fitzgerald Avenue bridge geotechnical investigation factual report. Prepared for Christchurch City Council. 2006.

[7] Diana-J3. Finite-element program for effective stress analysis of two-phase soil medium. Taisei Corporation internal report 1997 [in Japanese] ed: Software science, 1987.

[8] Bowen H and Cubrinovski M. Effective stress analysis of piles in liquefiable soil: A case study of a bridge foundation. *Bulletin of the New Zealand Society for Earthquake Engineering* 2008; **41**(4): 247-262.

[9] Cubrinovski M and Bradley BA. Evaluation of seismic performance of geotechnical structures, in *International Conference on Performance-Based Design in Earthquake Geotechnical Engineering — from case history to practice*, Tokyo, Japan, 2009, 16.

[10] Cubrinovski M and Ishihara K. Modelling of Sand Behaviour based on state concept. *Soils and Foundations* 1998; **28**(3): 115-127.

[11] Cubrinovski M and Ishihara K. State Concept and Modified Elastoplasticity for Sand Modelling. *Soils and foundations* 1998; **38**(4): 213-225.

[12] Ishihara K. Liquefaction and flow failure during earthquakes. *33rd Rankine Lecture, Geotechnique* 1993; **43**(3): 351-415.

[13] Cubrinovski M, Uzuoka R, Sugita H, Tokimatsu K, Sato M, Ishihara K, Tsukamoto Y, and Kamata T. Prediction of pile response to lateral spreading by 3-D soil-water coupled dynamic analysis: shaking in the direction of ground flow. *Soil Dynamics and Earthquake Engineering* 2008; **28**(6): 421-435, DOI: 10.1016/j.soildyn.2007.10.015.

[14] Kramer SL. *Geotechnical Earthquake Engineering*. Prentice-Hall: Upper Saddle River, NJ., 1996; 653.

[15] Cornell CA. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 1968; **58**(5): 1583–1606.

[16] Bradley BA, Cubrinovski M, Dhakal RP, and MacRae GA. Intensity measures for the seismic response of pile foundations. *Soil Dynamics and Earthquake Engineering* 2009; **29**(6): 1046-1058, DOI: 10.1016/j.soildyn.2008.12.002.

[17] Bradley BA, Cubrinovski M, MacRae GA, and Dhakal RP. Ground motion prediction equation for spectrum intensity from spectral acceleration relationships. *Bulletin of the Seismological Society of America* 2009; **99**(1): 277-285, DOI: doi: 10.1785/0120080044.

[18] Boore DM and Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s. *Earthquake Spectra* 2008; **24**(1): 99-138.

[19] Baker JW and Cornell CA. Spectral Shape, record selection and epsilon. *Earthquake Engineering and Structural Dynamics* 2006; **35**(9): 1077-1095.

[20] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. *Report No. RMS-35, RMS Program*,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[21] Stirling MW, Gerstenberger M, Litchfield N, McVerry GH, Smith WD, Pettinga JR, and Barnes P. Updated probabilistic seismic hazard assessment for the Canterbury region. 2007. 58.

[22] Power M, Chiou B, Abrahamson NA, Bozorgnia Y, Shantz T, and Roblee C. An Overview of the NGA Project. *Earthquake Spectra* 2008; **24**(1): 3-21.

[23] Luco N and Bazzurro P. Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses? *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1813-1835, DOI: 10.1002/eqe.695.

[24] Boore DM, Watson-Lamprey J, and Abrahamson NA. Orientation-Independent Measures of Ground Motion. *Bulletin of the Seismological Society of America* 2006; **96**(4A): 1502-1511, DOI: 10.1785/0120050209.

[25] Hopkins W. Cost estimates for seismic damage to the Fitzgerald Avenue twin bridges. 2009, (personal communication).

[26] MWH. Avonside/Fitzgerald/Kilmore Intersection: Economic Evaluation. Prepared for: Christchurch City Council, 2008. 42.

[27] Bradley BA and Lee DS. Correlations in structure-specific seismic loss estimation. *Earthquake Engineering and Structural Dynamics* 2009: (in press).

[28] Microsoft Project. http://en.wikipedia.org/wiki/Microsoft_Project. Last accessed: 21 March

[29] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[30] Krawinkler H. Van Nuys hotel building testbed report: exercising seismic performance assessment. PEER 2005/11 Pacific Earthquake Engineering Research Centre, 2005. 264.

[31] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: from structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2009: (submitted).

15.Conclusions and Recommendations

In this dissertation, probabilistic seismic risk assessment methods have been considered for use in the assessment of the seismic performance of specific structures. The use of such structure-specific probabilistic seismic risk assessments requires: the estimation of seismic hazard, structural response, component damage, component loss, and global loss measures for the entire structure; as well as a rigorous and computationally efficient probabilistic methodology to convolve these various tasks. This dissertation has made contributions in all facets of such a seismic performance assessment methodology mentioned above. The following sections summarize the key findings of this dissertation, its limitations, and potential future work.

15.1 Key contributions

15.1.1 Simplified methods in probabilistic seismic risk assessment

Chapters 3, 4, and 8 focused on the efficacy of various simplifications when conducting seismic risk assessments. Simplified methods of seismic risk analysis will always be desired, just as simplified methods of seismic response analysis (e.g. equivalent linear analysis or non-linear static analysis) are routinely used in conventional seismic design. However, in order for a simplified method to be successful it must provide the following three things: i) a reduction in the level of input information required to perform the analysis; ii) a reduction in the computational time to perform the analysis; and iii) a reduction in the level of expertise to set up the analysis, process and interpret the results.

In regard to the above three points it can be stated that while the power law seismic

hazard model discussed in Chapters 3 and 4 requires only knowledge of the parameters k, k_0, as illustrated in Chapter 4, these parameters should be varied as a function of level of demand or annual rate of exceedance considered. Therefore the analyst must have access to seismic hazard data to obtain the necessary parameters for a given region of the curve considered. The computational time for closed-form solutions using the power-law hazard model is less than solution via numerical methods, however as the numerical solution can be performed in the order of several seconds, this is not a significant benefit of the closed-form solution. Thirdly, while the initial development of such an algorithm to perform the numerical integration may require some effort, once this is completed there is no difference between the closed-form and numerical solutions in terms of user expertise to set up and perform the analysis. Thus based on the above discussion it should be clear that it is the authors' opinion that such use of these closed form solutions should be abandoned (apart from providing insight). What is really required is easy access for analysts to the results of state-of-the-art probabilistic seismic hazard analysis and simple software to perform the probabilistic risk calculations numerically.

Examination of the aforementioned three requirements of simplified methods in regard to the use of the FOSM method discussed in Chapter 8 reveals several points. Firstly, only the first two moments of the input variables are necessary to compute the FOSM approximation of the loss given intensity ($L|IM$) relationship, as opposed to the entire distribution necessary for direct numerical solution. However, since the lognormal distribution, which is uniquely defined by its first two moments, is commonly used as the probability distribution for input variables then there is no benefit in the required input information using the FOSM method. Computational demands discussed in Chapter 8 illustrate that computationally seismic risk assessments can be significantly demanding; particularly depending on the assumptions made regarding component correlations (Chapter 9 also illustrates this). Therefore, for an initial assessment, if the level of error in the FOSM approximation, as illustrated in Chapter 8, is deemed acceptable then there is a clear benefit in its use. Regarding the level of expertise to conduct and interpret the analysis, initially it is more demanding to develop the algorithms for the direct numerical integration solution than using the FOSM method, however once these algorithms have been developed the input information and output results of the two methods is identical.

15.1.2 Ground motion prediction of spectrum intensity

Chapters 5, 6, and 13 illustrated that the ground motion intensity measure: spectrum intensity, *SI*, is a desirable intensity measure for efficiently predicting the seismic response of spatially distributed accelerations and displacements in frame structures, and the seismic response of soil deposits and pile foundations embedded in such deposits. In order for *SI* to be used in probabilistic seismic risk assessments it is a prerequisite to have advanced ground motion prediction equations (GMPE's) so that the *SI* seismic hazard can be computed for the site under consideration.

Noting the *SI* is defined as the integral of the pseudo-spectral velocity (*PSV*) from 0.1 to 2.5 seconds, and that *PSV* is directly related to pseudo-spectral acceleration (S_a), Chapter 8 presents a method by which the first two moments (i.e. mean and variance) of *SI* can be obtained from GMPE's for S_a. Rigorous statistical distribution tests are used to illustrate the accuracy of the assumed lognormal distribution for *SI* based on the analytically derived statistical moments. It is illustrated that *SI* is an intensity measure with significantly lower uncertainty than spectral accelerations, due to the non-perfect (i.e. partial) correlation of spectral acceleration ordinates at different vibration periods.

This approach for predicting *SI*, as opposed to developing a regression equation based directly on empirical data, is advantageous in that GMPE's for S_a are the most advanced of any intensity measure. Thus, the current approach allows for both state-of-the-art prediction of *SI*, as well as the flexibility of using different S_a GMPE's for different tectonic regions.

15.1.3 Ground motion intensity measures and ground motion selection

Ground motion intensity measures (IM's) and input ground motion records provide the link between seismic hazard analysis and seismic response analysis. In the context of probabilistic seismic risk assessments, probabilistic seismic hazard analysis (PSHA) provides the temporal occurrence of various levels of the ground motion IM, while probabilistic seismic response analysis provides a probabilistic relationship between the ground motion IM and one or more measures of seismic response (so-called engineering demand parameters, EDP's). It is discussed in Chapters 5, 6 and 13 that key factors in the selection of a ground motion IM are *predictability*, *efficiency*, and *sufficiency*. *Predictability* and *efficiency* relate to the aleatory uncertainty in ground motion prediction and seismic response relationships, while *sufficiency* relates to the bias induced in the seismic response relationships due to differences

in the characteristics of the adopted ground motion records used in seismic response analyses with those that will occur at the site in the future.

Chapter 5 investigated the effect of intensity measure selection in the prediction of spatially distributed acceleration and displacement demands in a 10 storey frame structure. It was observed that conventional IM's such as S_a at the fundamental period of the structure efficiently predict displacement demands dominated by the first mode of vibration, but poorly predict displacement demands due to higher-modes and peak floor accelerations. Velocity-based IM's such as *SI* and peak ground velocity (*PGV*) were found to provide moderately efficient prediction of both displacement and acceleration demands. When the spatially distributed seismic response estimation using various ground motion IMs was combined with ground motion hazard, component fragility, and component loss functions, the seismic risk assessment results (e.g. annual rate of demand, probability of collapse, distribution of loss given IM, annual rate of direct repair loss) obtained in Chapter 6 illustrated that these velocity-based IMs were most appropriate for this structure. These observations were attributed to the aforementioned points as well as the fact that uncertainty in ground motion prediction of the IM (*predictability*) is also an important factor, and this was notably better for *SI* and *PGV* than other IMs considered.

Chapter 13 examined the performance of various ground motion IM's for seismic response analysis of pile foundations in liquefiable and non-liquefiable soils. It was found that velocity-based IM's (namely *SI* and *PGV*), which wave propagation theory suggests are related to peak shear strains, best predicted peak displacements at the surface of the soil deposits, as well as the peak displacement and curvature of the pile foundations embedded in such soils. As noted in the previous paragraph, *SI* and *PGV* are both IM's with good *predictability*, thus further promoting them as IM's for the seismic response of soil deposits and pile foundations.

It was observed in Chapters 5, 6 and 13 that there is a clear trend between the efficiency and sufficiency of an IM for predicting seismic response. Thus an IM which is efficient at predicting a specific EDP will also tend to be relatively sufficient also. Chapter 5 however illustrated that for a 10 storey frame structure, no common IM's were sufficient (i.e. unbiased) in predicting spatially distributed displacement and acceleration demands. The result of this finding in Chapter 5 is that pragmatic estimation of seismic response will always have some dependence on ground motions used in seismic response analyses. Two possible options to account for the importance of ground motion selection were offered in this dissertation. Firstly, in Chapter 14, where seismic response was assessed for 9 different values of *PGA*,

seismic hazard deaggregation was used to determine the causal magnitude, distance and epsilon values of contributing ground motions to the seismic hazard. These deaggregation statistics, as well as a requirement of an amplitude scale factor between 0.6 – 1.6 were used to select ten two-component ground motions for seismic response analyses of the Fitzgerald Avenue twin bridges at each of the 9 intensity values of interest. Secondly, Chapter 6 proposed a method, based on the regression parameters obtained from sufficiency analyses, by which the statistics of obtained seismic response analyses could be corrected for differences in the characteristics of input ground motions used compared to those which are deemed to be most likely to occur at the site in the future.

15.1.4 Efficient solution of probabilistic seismic risk assessments

The computational times required in performing seismic risk assessments presented in Chapters 8 and 9 illustrated that consideration should be given to the computational and numerical details of such assessments to ensure that such demands are not considered a restriction of using such methods. Given that the solution of numerous integral equations represents the primary computational demand in such assessments, Chapter 7 developed an efficient numerical integration algorithm for their evaluation. The algorithm, termed Magnitude-oriented Adaptive Quadrature (MAQ), was developed to be an algorithm with both locally and globally adaptive capabilities, making it significantly more efficient than conventional integration algorithms.

Attention was also devoted to the development of methodological frameworks which were amenable to efficient numerical solution. In particular, Chapter 9 proposed a numerically efficient method by which the effect of correlations between fragility function damage states of different components can be accounted for. The method uses the standard bi-variate normal cumulative density function (CDF) for which highly efficient computational methods exist, as opposed to previous solutions via Monte Carlo simulation, or a less general solution via optimisation.

The large computational cost of direct solution of the seismic risk assessment equations via numerical integration was also a primary motivator in the exploration of the efficacy for using approximate moment-estimation methods in Chapter 8, such as the first-order second-moment method.

While we are currently in an age of exponentially increasing computational resources, the increasing expectations of users, and increasing computational complexities which will be

manifested by an increase in the sophistication of seismic risk assessment methodologies and their application will ensure that computational aspects should be borne in mind in the future development of such methodologies.

15.1.5 Component correlations in seismic risk assessment

In addition to uncertainties in the *EDP|IM*, *DS|EDP* and *L|DS* relationships for a single component, their also exist statistical dependencies, or correlations, between the multiple components which comprise a structure. Chapter 9 was concerned with the consideration of such component correlations in seismic loss estimation. The consideration of component correlations by previous researchers has been limited by methodological tractability, increased computational demand, and a paucity of data for their computation (e.g. component correlations were neglected in Chapter 2). Chapter 9 presented a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered. Methods to determine the necessary correlations were discussed, with particular emphasis on those which can be used in the absence of sufficient empirical data and rely somewhat on judgement. Methods of accounting for epistemic uncertainties in the estimation of correlation coefficients were also explored.

The effects of various assumptions regarding correlations were illustrated via application to a case-study office structure. The observation that certain correlation assumptions can lead to errors in excess of 50% in the lognormal standard deviation in the loss given intensity and loss hazard relationships clearly indicates the importance of their careful consideration. Consideration of partial component correlations was also observed to increase the computational demands of performing seismic risk assessments with full consideration of partial correlations requiring in excess of 50-times more computational time than other correlation assumptions for the example examined.

15.1.6 Epistemic uncertainties in seismic risk assessment

All aspects in the estimation of seismic risk of a specific structure involve some degree of uncertainty. The PEER framework formula adopted throughout this dissertation allows for explicit consideration of apparently random realisations in each of aspect of seismic risk assessment via couching the seismic risk problem in a fully probabilistic framework. In addition to uncertainty due to apparent randomness in each aspect of seismic risk, there also

exists potentially significant uncertainty due to a lack-of-knowledge in the prediction of such aspects. In earthquake engineering circles, these two distinctly different natured uncertainties are commonly referred to as aleatory and epistemic uncertainties, respectively.

Epistemic uncertainties have been considered in probabilistic seismic hazard analysis (PSHA) for over 20 years, however new details of these uncertainties are still emerging as scientists make further advances via empirical observations and theoretical developments. Chapter 10 investigated epistemic uncertainty trends in the PSHA results for various sites and IM's in the San Francisco bay area. The study used the Working Group on California Earthquake Probabilities (WGCEP) 2002 earthquake rupture forecast (ERF), which is widely regarded as the most advanced ERF ever developed. State-of-the-art ground motion prediction equations developed as part of the Next Generation Attenuation (NGA) project, as well as several other predecessor equations were adopted to consider epistemic uncertainties in ground motion prediction. It was found that for the particular sites and ground motion IM's considered there was a clear trend for increasing epistemic uncertainty with reducing annual probability of exceedance. The absolute magnitude of the epistemic uncertainty was also (somewhat obviously) sensitive to the alternative ground motions used. In certain cases the empirical distribution of the annual probability of exceedance for a given value of the IM was found to be well represented by the lognormal distribution. Examination of the correlation structure of the epistemic uncertainty in the annual probability of exceedance illustrated that it is a function of the ratio of the two IM values considered, and a simple predictive equation was suggested in the absence of more detailed information.

Despite the generally accepted importance of epistemic uncertainties in PSHA; conventional seismic risk assessments, with few exceptions, consider only the mean seismic hazard curve and do not explicitly account for, or separately propagate, epistemic uncertainties. Chapter 10 illustrated and compared three methods by which epistemic uncertainties in PSHA can be propagated to other measures of seismic risk, such as the annual probability of structural collapse.

Chapter 11 examined epistemic uncertainties in component fragility functions. Currently fragility functions are developed primarily using solely experimental, numerical or observational data, and therefore only capture aleatory uncertainty. Various epistemic uncertainties in the development of component fragility functions were identified. Methods for the consideration, combination and propagation of such epistemic uncertainties were discussed. In particular, epistemic uncertainty associated with using finite datasets was illustrated via application to two components whose fragility functions have been published in

technical literature.

Chapter 9 also discussed methods for the consideration of epistemic uncertainty in component correlations used in seismic loss estimation. Given that assumptions made regarding component correlations can have a significant effect on the results of seismic risk assessments, and that these assumptions are commonly made in the absence of numerous high-quality empirical observations, it is prudent to explicitly consider such epistemic uncertainties.

15.1.7 State-of-the-art applications of seismic risk assessment methods

In addition to the development of seismic risk assessment methodologies that form the core of this dissertation, attention has also been given to the application of such methodologies to actual structures. The two structures considered were: i) a typical 10 storey reinforced concrete frame structure which contains typical commercial densities of non-structural components and contents (Chapter 2); and ii) the Fitzgerald Avenue twin bridges, which are located along one of four arterial avenues in Christchurch, New Zealand (Chapter 14).

In addition to the two aforementioned structures providing a platform for illustrating the state-of-the-art in structure-specific seismic risk assessment methods, several additional novel contributions were present in each of these applications. Firstly, the seismic risk assessment of the typical New Zealand office structure included contents in addition to the structural and non-structural components previously considered by others. It was found that contents such as computers, server and network equipment provide a significant source of seismic risk in a system-level assessment. Secondly, the seismic risk assessment of the Fitzgerald Avenue twin bridges was the first structure-specific seismic risk assessment of a lifeline infrastructure considering both direct repair costs and loss of functionality costs.

In the seismic risk assessment of the Fitzgerald Avenue twin bridges (Chapter 14), emphasis was also given to all of the preliminary steps which should be given due consideration before arriving at final seismic risk assessment performance metrics. In particular, emphasis was placed on the importance of careful selection of input ground motions which match the results of probabilistic seismic hazard analysis deaggregation, as well as the importance of examining in great detail the seismic response of the system for a single input ground motion. The latter point in particular is important, because it is common in many aspects of engineering to witness confrontations as to the superiority of deterministic

vs. probabilistic assessments. The author wishes to emphasise that in terms of seismic response analysis, it is considered absolutely profound that an analyst first rigorously examines the seismic response of (in the case of Chapter 14) the entire soil-foundation-structure system both for the purposes of model verification and validation, but also to confirm *a priori* expectations or correct erroneous ones. This step is effectively deterministically examining the response of the system, and for the reasons noted above, is considered by the author as imperative before engineering demand parameters are assigned and the seismic response for multiple ground motions at multiple levels of intensity is considered (i.e. as part of a probabilistic risk analysis).

15.2 Limitations and future work

15.2.1 Intensity measures for seismic risk assessment

Chapters 5 and 6 investigated the effect of ground motion intensity measure (IM) selection for the seismic response and loss assessment of a 10 storey office structure, while Chapter 13 investigated the effect of IM selection on the seismic response assessment of pile foundations. As mentioned in these chapters, there has been significant interest by past researchers on ground motion intensity measures for the determination of the peak interstorey drift ratio over all floors in frame-structures. The interest in peak interstorey drift ratio over all floors is that it related well to global instability of the structure. The consideration of spatially distributed demands when selecting IMs is important in the seismic loss estimation of a structure because of the spatially distributed nature of the components used in the loss estimation methodology and also that components are sensitive to both displacements and accelerations.

In Chapters 5, 6 and 13 it was found that among the candidate IMs considered, spectrum intensity, *SI*, was found to be the 'optimal' IM. This conclusion is however based on the seismic response analyses of a 10 storey structure (Chapters 5 and 6) and a simplified soil-pile-structure model (Chapter 13), and should be empirically examined for other types of structural and geotechnical systems. In particular, the effects of intensity measure selection for short-period structures may yield different results than for the longer period 10 storey structure considered in Chapters 5 and 6.

Because of the importance of peak accelerations (high frequency) and peak

displacements (lower frequency) in the seismic loss estimation of the 10 storey structure considered in Chapter 6 it may be beneficial to investigate the use of a vector IM (e.g. [1]) comprised of high and low frequency IMs. The vector IM approach may also be useful for seismic effective stress analyses in which the demands on a pile foundation are dependent on both the kinematic soil displacements as well as the superstructure response.

15.2.2 Ground motion selection

As long as pragmatic intensity measures are used, which will be imperfect predictors of seismic demands, it will be important to carefully consider the ground motions which are considered in seismic response analyses. In order to facilitate appropriate ground motion selection, seismic hazard curves and their deaggregation for various levels of intensity measure/rate of exceedance must be easily accessible to users. The use of deaggregation allows analysts to understand the characteristics of the earthquake sources which most significantly contribute to the seismic hazard at the site, and therefore those ground motions which are most representative for use in seismic response analysis.

In addition to the point above regarding ground motion selection, seismic risk analysts must realize the errors which are present from excessively scaling and/or manipulating ground motion records (such as when conducting incremental dynamic analysis (IDA) or excessively modifying the frequency content of ground motions to 'match' some target response spectra). The use of physics-based earthquake rupture and ground motion simulations will aid in the task of ground motion selection, since simulated ground motions for each specific site of interest will be used to construct the seismic hazard curve, and then can also be used directly in seismic response analyses. Thus un-scaled ground motions could be obtained which are consistent with the source scaling properties of earthquakes. Unfortunately, it is likely to be several years before such physics –based simulations extend beyond their current verification and validation stages, particularly for high frequency waves which require small numerical grids for low velocity regions, and are currently simulated stochastically using ray theory.

15.2.3 Improved engineering demand parameters

In all probabilistic seismic performance assessments that the author is aware of the engineering demand parameter (EDP) used is some peak quantity, such as peak interstorey drift for displacement-sensitive components, or peak floor acceleration for acceleration-

sensitive components. As discussed to some extent in Chapter 11 there is a need to develop EDP's that consider more completely the seismic demand of structures, namely peak response, number of cycles, cumulative effects etc. As noted in Chapter 11, the lack of attention that this topic has received relative to the topic of optimal intensity measures is likely due to the fact that data for such EDP's is needed when developing component fragility functions. There are however clear benefits in using such advanced EDP's and despite the initial complication of developing fragility functions for them, the author expects they will be adopted in future. Once such EDP's become considered it will also be necessary to re-examine the efficiency and sufficiency of the predominantly used ground motion IMs (e.g. *PGA*, S_a, *SI*, *PGV*). If such EDP's consider cumulative effects then clearly the duration of strong ground motion shaking must be an important factor. Seismic hazard deaggregation as discussed in the previous section does not currently provide estimates of the expected durations of ground motion shaking and also duration, while being considered important (e.g. [2]), is not yet unanimously advocated in ground motion selection (e.g. [3]).

15.2.4 Consideration of human injuries

Consequences of seismic damage to structural and geotechnical systems in this dissertation were measured in terms of direct repair cost as well as loss of functionality. An obvious omission was the consequences due to human injuries (both fatal and non-fatal). Current quantitative methods of estimating human injuries (e.g. [4]) consider only human fatalities due to structural collapse. However, Porter *et al*. [5] illustrated that for the Northridge earthquake, 97% of the economic consequence of injury costs were due to non-fatal injuries.

The consideration of a probabilistic structure-specific methodology to account for various levels of occupant injury is confronted by two difficulties. Firstly, consideration of a structure-specific method means, for example, accounting for occupant injuries due to falling contents. The number of different considerations leads to not only complexity in terms of numerical book-keeping but, more importantly; high computational costs in the consideration of correlations between multiple contents etc. (e.g. see Chapter 9 for the effect of correlations in computational demands). Secondly, confusion exists as to the use of such an occupant injury model, given that the primary stakeholder of interest would be local authorities (interested in non-fatal injuries at a regional level to assess demands on hospitals etc.). Furthermore, there is an incompatibility of having a building-specific injury model with a

region-wide (and therefore less detailed) inventory model.

15.2.5 Loss amplification and the interaction between direct repair cost and duration of loss of functionality

Loss amplification reflects all the ways in which the costs incurred for a certain level of damage become amplified when the damage occurs as part of a major catastrophe. Boissonnade *et al*. [6] provide discussion on the recent experiences of loss amplification in the US (due to primarily hurricane catastrophes). The core components of loss amplification include the commonly referred term "demand surge" (an increase in unit cost due to an increased demand and limited capacity) as well as repair cost delay inflation, claims inflation and coverage expansion. Throughout this dissertation the effects of loss amplification have been purposefully neglected, because sound models to predict such phenomena are not yet developed.

There is also an important interaction between direct repair cost, repair duration, and loss amplification, in that damaged assets for which stakeholders are prepared to pay significantly inflated costs to repair, will be repaired more swiftly, than damaged assets whose stakeholders are not prepared to pay inflated costs. The relationship between increased repair cost – reduced repair duration, (or vice-versa) will clearly be specific to various regions of the world. As such, regional catastrophe models are likely needed to determine the level of regional damage which will result from various earthquake scenarios, which can then be used to aid in the consideration of loss amplification for building-specific risk assessments such as those this dissertation is concerned with.

15.2.6 Consideration of aftershock ground motions

In addition to seismic damage to structures due to main-shock induced ground motions, significant risk may also be present due to aftershock ground motion hazards. A preliminary body of work on hazard and risk assessments considering aftershocks is given by Yeo and Cornell [7]. Consideration of the risk of aftershock ground motions poses several complications to current seismic risk assessment methodologies such as those discussed in this dissertation. Firstly, the development of aftershock probabilistic seismic hazard analysis (APSHA) depends on the characteristics of the main-shock, and is time-dependent. Secondly, consideration of the seismic response of structures under main-shock aftershock sequences

involves back-to-back seismic response analyses which will be more computationally demanding than seismic response analyses for main-shock consideration only. Thirdly, computation of seismic risk with explicit consideration of aftershocks requires the use of renewal models for simulating the seismic risk, which is significantly more complicated and computationally demanding than current methods which use integration with the seismic hazard curve.

Because of the aforementioned difficulties in the consideration of aftershocks it is pertinent that their significance be assessed rigorously via several case study structures (examples presented in Yeo and Cornell [7] while useful are, in the authors' opinion, too simple to provide a significant level of insight).

15.2.7 Explicit epistemic uncertainty consideration and propagation in seismic loss estimation

Explicit consideration of epistemic uncertainties is common-place in probabilistic seismic hazard analysis and it would seem obvious that this will, over time, extend to all of the various relationships in seismic risk assessment methods. This thesis has illustrated some features of epistemic uncertainties in seismic hazard analyses (Chapter 10) and component fragility functions (Chapter 11) and how they can be propagated in seismic risk assessments. There are also various studies investigating the epistemic uncertainties in the seismic response of structures (e.g. [8, 9]), while no literature, to the authors knowledge, is available for epistemic uncertainties in component loss functions.

A common conclusion of studies investigating epistemic uncertainties in the relationship between seismic response and seismic intensity is that the epistemic uncertainty is notably less than the aleatory uncertainty due to record-to-record variability (i.e. multiple input ground motions) [9, 10]. However, it is argued by the author that this view (and observations upon which it is based) are biased because: (i) improper selection of ground motion intensity measures and/or ground motion records leads to an increase in record-to-record variability; and (ii) the majority of researchers consider only uncertainty in the constitutive model parameters of an analytical model and not other epistemic uncertainties such as soil-structure interaction; 3-dimensional effects; uncertainty geometry (particularly soil layering); different constitutive models; effects of human error in construction; influence of non-structural elements; viscous nature and magnitude of damping. In the authors opinion, once the significance of these biases are reduced, epistemic and aleatory uncertainties are

likely to be of an equal significance, and epistemic uncertainties would be expected to dominate in situations of poorly known sites, structures, components, and consequences.

A significant amount of research is needed to develop epistemic uncertainty estimates for fragility and loss functions. This will however be a beneficial process in that it will enable the identification of those components whose uncertainty causes a significant uncertainty on the seismic risk of the system considered. While epistemic uncertainties in seismic hazard analyses have been considered for over 30 years [11], the lack of availability of such results in the public domain has limited their explicit use to-date.

15.2.8 Efficient numerical solution of the PBEE equations

It was illustrated that seismic risk assessment computations for structures with many components and consideration of correlations (Chapters 8 and 9) can be very computationally demanding. If in addition to the seismic risk assessment methodologies considered in this dissertation, additionally important factors such as: human injuries, loss amplification, aftershocks, and explicit epistemic uncertainties are also considered then one can imagine the computational demands will require a greater emphasis on computational algorithms. Chapter 7 of this thesis was concerned with the development of an efficient numerical integration algorithm which forms a key part of the numerical solution of the seismic risk assessment equations. Future work could also focus on the implementation of seismic risk assessment codes for parallel-processing, something which clearly the solution procedures are very amenable to. As with all software, in addition to computational efficiency, graphical user interfaces (GUI's) will aid in the uptake of such software by practitioners, however in the author's opinion at the present time this should be of secondary concern to developing such methods further.

15.2.9 Case study applications using rigorous seismic risk assessments

Chapters 2 and 14 were concerned with illustrating the application of rigorous seismic risk assessment methods for two case study structures. Mention was also given in these chapters to other similar recent case-studies. Despite these studies, there still remains a clear need for further illustration of the benefits to stakeholders of using such methodologies. Such case-studies could be used to benchmark the seismic performance of code-conforming and earthquake-prone structures, as well as the potential benefits of emerging seismic

technologies.

Similar to the development of probabilistic seismic hazard analysis, probabilistic seismic risk assessment frameworks will greatly benefit by active participation of end-users. Stakeholders, both from the public and private sector, have different needs regarding the outputs of seismic risk assessments and effective communication of these needs to framework developers will lead to both a better end product and faster update of the methods in practice.

15.3 Concluding remarks

The summaries, observations and future recommendations made in this chapter are based on the results of the former chapters in this dissertation. In particular, while the majority of the key contributions in this dissertation were based on theoretical developments in structure-specific seismic risk assessment methodologies, examples and case-studies used to illustrate and validate such theoretical ideals represent a small subset of the numerous possibilities for which such methodologies could be used for in future. The interpretation of the conclusions in this dissertation should be made while keeping these limitations in mind. Given the insight that such structure-specific seismic risk assessment methods offer to stakeholders, further research into their development and implementation in the commercial environment appears warranted.

15.4 References

[1] Baker JW. Vector-valued Ground Motion Intensity Measures for Probabilistic Seismic Demand Analysis. Ph.D Thesis, Department of Civil and Environmental Engineering Stanford University, 2005, 347.

[2] Bommer JJ, Magenes G, Hancock J, and Penazzo P. The influence of string-motion duration on the seismic response of masonary structures. *Bulletin of Earthquake Engineering* 2004; **2**(1-26.

[3] Baker JW and Cornell CA. Spectral Shape, record selection and epsilon. *Earthquake Engineering and Structural Dynamics* 2006; **35**(9): 1077-1095.

[4] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[5] Porter KA, Shoaf K, and Seligson H. Value of injuries in the Northridge earthquake. *Earthquake Spectra* 2006; **22**(2): 555-563, DOI: DOI: 10.1193/1.2194529.

[6] Boissonnade A. Modelling demand surge. *Workshop on Risk Acceptance and Risk Communication*, Stanford University, CA, 2007.

[7] Yeo GL and Cornell CA. Stochastic Characterisation and Decision Bases under Time-Dependent Aftershock Risk in Performance-based Earthquake Engineering. John A. Blume Earthquake Engineering Centre, Stanford University., CA, 2005.

[8] Liel A, Haselton CB, Deierlein GG, and Baker JW. Incorporating modelling uncertainties in the assessment of seismic collapse risk of buildings. *Structural Safety* 2009; **31**(2): 197-211, DOI: 10.1016/j.strusafe.2008.06.002.

[9] Dolsek M. Incremental dynamic analysis with consideration of modelling uncertainties. *Earthquake Engineering and Structural Dynamics* 2009, DOI: 10.1002/eqe.869.

[10] Mehanny SSF and Ayoub AS. Variability in inelastic displacement demands: uncertainty in system parameters versus randomness in ground records. *Engineering Structures* 2008; **30**(1002-1013., DOI: 10.1016/j.engstruct.2007.06.009.

[11] Kulkarni RB, Youngs RR, and Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. *8th World Conference on Earthquake Engineering*,, San Francisco, CA, 1984.

Appendix A: Closed-form Solutions for Seismic Demand and Collapse Hazards

Below the closed form solutions for annual frequency of exceeding a specified demand and annual frequency of structural collapse are presented. The first closed-form solutions were published for the demand hazard in References [1, 2], and using similar assumptions, annual frequencies of limit state exceedance and structural collapse can also be computed [3-5].

A.1. Demand hazard

The annual frequency of exceeding a specified level of demand, $\lambda_{EDP}(edp)$, is given by:

$$\lambda_{EDP}(edp) = \int G_{EDP|IM}(edp|im)\left|\frac{d\lambda_{IM}(im)}{dIM}\right| dIM \tag{A-1}$$

where $G_{EDP|IM}(edp|im)$ is the complementary cumulative distribution function (CCDF) of *EDP* given *IM* = *im* and $\lambda_{IM}(im)$ is the annual frequency of exceeding *IM* = *im* (the seismic hazard curve).

Via integration by parts, Equation (A-1) becomes:

$$\begin{aligned}\lambda_{EDP}(edp) &= G_{EDP|IM}(edp|im)\lambda_{IM}(im)\Big]_0^{\infty} + \int \lambda_{IM}(im)\left|\frac{dG_{EDP|IM}(edp|im)}{dIM}\right| dIM \\ &= \int \lambda_{IM}(im)\left|\frac{dG_{EDP|IM}(edp|im)}{dIM}\right| dIM\end{aligned} \tag{A-2}$$

where the first term vanishes at both integration limits.

In order to obtain a closed-form solution specific functional forms for the ground motion hazard and *EDP*|*IM* relationship must be assumed. The seismic hazard is assumed to

be linear in log-log space, i.e.:

$$\lambda_{IM}(im) = k_0 IM^{-k} \tag{A-3}$$

where k and k_0 are parameters which are calibrated based on fitting Equation (A-3) to seismic hazard data obtained from probabilistic seismic hazard analysis (PSHA). The *EDP*|*IM* random variable is assumed to have a lognormal distribution (i.e. ln*EDP*|*IM* has a normal distribution); the mean value of ln*EDP*|*IM* is also a linear function of *IM*; and the variance of ln*EDP*|*IM* is independent of *IM*, i.e.:

$$\begin{aligned} &\ln EDP \Big| IM \sim N(\mu_{\ln EDP|IM}, \sigma^2_{\ln EDP|IM}) \\ &\mu_{\ln EDP|IM} = \ln a + b \ln IM \\ &\sigma^2_{\ln EDP|IM} = \sigma^2_R \end{aligned} \tag{A-4}$$

where a and b are parameters which are calibrated based on fitting Equation (A-4) to seismic response analysis data. Not that as *EDP*|*IM* is lognormally distributed, the exponent of the mean of ln*EDP*|*IM* is equal to the median of *EDP*|*IM* (i.e. $EDP\big|IM_{50} = aIM^b$).

From the definition of the CCDF, $G_{EDP|IM}(edp|im)$ can be expressed as:

$$\begin{aligned} G_{EDP|IM}(edp|im) &= 1 - F_{EDP|IM}(edp|im) \\ &= 1 - \Phi\left\{\frac{\ln edp - \mu_{\ln EDP|IM}}{\sigma^2_{\ln EDP|IM}}\right\} \\ &= 1 - \Phi\left\{\frac{\ln(edp/aIM^b)}{\sigma_R}\right\} \end{aligned} \tag{A-5}$$

where $\Phi\{\ \}$ is the cumulative normal variate. By making the substitution $u = \dfrac{\ln(edp/aIM^b)}{\sigma_R}$, the derivative of $G_{EDP|IM}(edp|im)$ can be computed as:

$$\begin{aligned} \frac{dG_{EDP|IM}(edp|im)}{dIM} &= \frac{dG_{EDP|IM}(edp|im)}{du}\frac{du}{dIM} \\ &= -\phi\{u\}.-\frac{b}{IM.\sigma_R} \end{aligned} \tag{A-6}$$

where $\phi\{\ \}$ is the normal variate (not to be confused with the cumulative normal variate

$\Phi\{\ \}$). Now the normal variate is given by:

$$\phi\{u\} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \tag{A-7}$$

By using Equations (A-2)-(A-7), the demand hazard is manipulated in the steps below:

$$\begin{aligned}
\lambda_{EDP}(edp) &= \int k_0 IM^{-k} \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left\{\frac{\ln edp - \ln a - b\ln IM}{\sigma_R}\right\}^2\right) dIM \\
&= k_0 \int \exp\left\{-\frac{1}{2}2k\ln IM\right\} \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left\{\frac{\ln\left(\frac{edp}{a}\right)^{1/b} - \ln IM}{\sigma_R/b}\right\}^2\right) dIM \\
&= k_0 \int \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left\{\frac{\left(\ln\left(\frac{edp}{a}\right)^{1/b} - \ln IM\right)^2}{\sigma_R^2/b^2} + 2k\ln IM\right\}\right) dIM \\
&= k_0 \int \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{b^2}{\sigma_R^2}\left\{\begin{array}{l}\left[\ln IM - \left(\ln\left(\frac{edp}{a}\right)^{1/b} - \frac{k\sigma_R^2}{b^2}\right)\right]^2 \\ -\frac{k^2\sigma_R^4}{b^4} + 2\frac{k\sigma_R^2}{b^2}\ln\left(\frac{edp}{a}\right)^{1/b}\end{array}\right\}\right) dIM
\end{aligned} \tag{A-8}$$

Making the substitution $v = \dfrac{\ln IM - \left(\ln\left(\frac{edp}{a}\right)^{1/b} - \frac{k\sigma_R^2}{b^2}\right)}{\sigma_R/b}$ gives $\dfrac{dIM}{dv} = \dfrac{IM.\sigma_R}{b}$ and hence Equation (A-8) simplifies further to:

$$\lambda_{EDP}(edp) = k_0 \exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2 - k\ln\left(\frac{edp}{a}\right)^{1/b}\right) \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}v^2\right) \frac{dIM}{dv} dv \tag{A-9}$$

The term within the integral of Equation (A-9) is simply the normal cumulative distribution function (CDF) and hence is equal to unity, thus giving the result:

$$\lambda_{EDP}(edp) = k_0 \exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2 - k\ln\left(\frac{edp}{a}\right)^{1/b}\right)$$
$$= k_0 \exp\left(-k\ln\left(\frac{edp}{a}\right)^{1/b}\right)\exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2\right) \quad \text{(A-10)}$$
$$\lambda_{EDP}(edp) = k_0\left(\frac{edp}{a}\right)^{-k/b}\exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2\right)$$

A.2. Demand hazard with epistemic uncertainty

If it is now considered that due to various sources epistemic uncertainty exists in the seismic hazard, then:

$$\tilde{\lambda}_{IM}(im) = k_0 IM^{-k}\tilde{\varepsilon}_{IM} \quad \text{(A-11)}$$

where $\tilde{\varepsilon}_{IM} \sim LN\left(\ln\left(1-0.5\sigma_{UIM}^2\right), \sigma_{UIM}^2\right)$, i.e. $\tilde{\varepsilon}_{IM}$ has a lognormal distribution with a mean of 1 and lognormal standard deviation σ_{UIM}^2. Similarly, epistemic uncertainty in the mean of the ln*EDP*|*IM* relationship is given by:

$$EDP|IM \sim LN(\mu_{\ln EDP|IM}, \sigma^2_{\ln EDP|IM}) = LN(\ln\left(aIM^b\tilde{\varepsilon}_{EDP}\right), \sigma_R^2) \quad \text{(A-12)}$$

where $\tilde{\varepsilon}_{EDP} \sim LN\left(\ln(1), \sigma_{UEDP}^2\right)$.

If $\tilde{\lambda}_{IM}(im)$ and $EDP|IM$ are considered conditioned on fixed (i.e. constant) values of $\tilde{\varepsilon}_{IM}$ and $\tilde{\varepsilon}_{EDP}$ then the parameters k_0 and a can be grouped with $\tilde{\varepsilon}_{IM}$ and $\tilde{\varepsilon}_{EDP}$, respectively. Thus the closed form solution with only aleatory uncertainty (Equation (A-10)) gives:

$$\lambda_{EDP|\varepsilon_{IM},\varepsilon_{EDP}}(edp|\varepsilon_{IM},\varepsilon_{EDP}) = k_0\left(\frac{edp}{a\varepsilon_{EDP}}\right)^{-k/b}\varepsilon_{IM}\exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2\right)$$
$$= k_0\left(\frac{edp}{a}\right)^{-k/b}\exp\left(\frac{1}{2}\frac{k^2}{b^2}\sigma_R^2\right)\left[\varepsilon_{EDP}^{k/b}\varepsilon_{IM}\right] \quad \text{(A-13)}$$

Now as the terms in square brackets contain the product of lognormal random variables raised to powers, it follows that $\lambda_{EDP}(edp)$ is also a lognormal random variable with mean and standard deviation given by:

$$E\left[\lambda_{EDP}(edp)\right] = k_0\left(\frac{edp}{a}\right)^{-k/b}\exp\left(\frac{1}{2}\frac{k^2}{b^2}\left(\sigma_R^2+\sigma_{UEDP}^2\right)\right)$$
$$Var\left[\ln\lambda_{EDP}(edp)\right] = \frac{k^2}{b^2}\sigma_{UEDP}^2+\sigma_{UIM}^2 \quad \text{(A-14)}$$

A.3. Collapse hazard

The annual frequency of structural collapse, λ_C, is given by:

$$\lambda_C = \int P_{C|IM}(im)\left|\frac{d\lambda_{IM}(im)}{dIM}\right|dIM \quad \text{(A-15)}$$

where $P_{C|IM}(im)$ is the probability of collapse given *IM* = *im* and $\lambda_{IM}(im)$ is the annual frequency of exceeding *IM* = *im*, as in the case of the demand hazard discussed above.

Similarly, to Equation (A-2) Integration by parts of Equation (A-15) yields:

$$\lambda_C = \int \lambda_{IM}(im)\left|\frac{dP_{C|IM}(im)}{dIM}\right|dIM \quad \text{(A-16)}$$

The seismic hazard curve is assumed to linear in log-log space as in Equation (A-11), while IM value causing structural collapse is assumed to have a lognormal distribution with median η and lognormal variance σ_{RC}^2, i.e.

$$IM|C \sim LN(\mu_{\ln IM|C}, \sigma_{IM|C}^2) = LN(\ln(\eta), \sigma_{RC}^2) \quad \text{(A-17)}$$

Hence the probability of collapse, $P_{C|IM}(im)$ and its derivative are given by:

$$P_{C|IM}(im) = \Phi\left\{\frac{\ln(im/\eta)}{\sigma_{RC}}\right\}$$
$$\frac{dP_{C|IM}(im)}{dIM} = \frac{1}{IM.\sigma_{RC}}\phi\left\{\frac{\ln(im/\eta)}{\sigma_{RC}}\right\} \quad \text{(A-18)}$$

where $\Phi\{\ \}$ is the cumulative normal variate and $\phi\{\ \}$ is the normal variate.

Using Equations (A-7) and (A-16)-(A-18), the collapse hazard can be manipulated as:

$$
\begin{aligned}
\lambda_C &= \int k_0 IM^{-k} \frac{1}{IM.\sigma_{RC}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left\{ \frac{\ln im - \ln \eta}{\sigma_{RC}} \right\}^2 \right) dIM \\
&= k_0 \int \exp\left\{ -\frac{1}{2} 2k \ln IM \right\} \frac{1}{IM.\sigma_{RC}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_{RC}^2} \{\ln im - \ln \eta\}^2 \right) dIM \\
&= k_0 \int \frac{1}{IM.\sigma_{RC}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_{RC}^2} \{\ln im - \ln \eta\}^2 - 2k \ln IM \right) dIM \\
&= k_0 \int \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_{RC}^2} \left\{ \begin{array}{l} \left[\ln im - \left(\ln \eta - 2k\sigma_{RC}^2\right)\right]^2 \\ -\left(k\sigma_{RC}^2\right)^2 + 2k\sigma_{RC}^2 \ln \eta \end{array} \right\} \right) dIM \\
&= k_0 \int \frac{b}{IM.\sigma_R} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_{RC}^2} \left\{ \left[\ln im - \left(\ln \eta - 2k\sigma_{RC}^2\right)\right]^2 \right\} \right) \text{x} \\
&\quad \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2 - k \ln \eta \right) dIM
\end{aligned}
\tag{A-19}
$$

Making the substitution $v = \dfrac{\ln im - \left(\ln \eta - 2k\sigma_{RC}^2\right)}{\sigma_{RC}}$ gives $\dfrac{dIM}{dv} = IM.\sigma_{RC}$ and hence Equation (A-19) simplifies further to:

$$
\lambda_C = k_0 \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2 - k \ln \eta \right) \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} v^2 \right) \frac{dIM}{dv} dv \tag{A-20}
$$

The term within the integral of Equation (A-20) is simply the normal cumulative distribution function (CDF) and hence is equal to unity, thus giving the result:

$$
\begin{aligned}
\lambda_C &= k_0 \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2 - k \ln \eta \right) \\
\lambda_C &= k_0 \eta^{-k} \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2 \right)
\end{aligned}
\tag{A-21}
$$

A.4. Collapse hazard with epistemic uncertainty

As in the case of the demand hazard, epistemic uncertainty in the seismic hazard is given by Equation (A-11), while the median IM value at which collapse occurs, η, is assumed to be random with a lognormal distribution, i.e.:

$$
IM|C \sim\sim LN(\mu_{\ln IM|C}, \sigma^2_{\ln IM|C}) = LN(\ln(\eta \tilde{\varepsilon}_C), \sigma^2_{RC}) \tag{A-22}
$$

where $\tilde{\varepsilon}_C \sim LN(\ln(1), \sigma^2_{UC})$.

As in the case of the demand hazard with epistemic uncertainty, the collapse hazard conditioned on fixed $\tilde{\varepsilon}_{IM}$ and $\tilde{\varepsilon}_{C}$ values is given by (Equation (A-21)):

$$\begin{aligned} \lambda_{C|\varepsilon_{IM},\varepsilon_C} &= k_0 \left(\eta \varepsilon_C\right)^{-k} \varepsilon_{IM} \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2\right) \\ &= k_0 \eta^{-k} \exp\left(\frac{1}{2} k^2 \sigma_{RC}^2\right) \left[\varepsilon_C{}^{-k} \varepsilon_{IM}\right] \end{aligned} \tag{A-23}$$

Now as the terms in square brackets contain the product of lognormal random variables raised to powers, it follows that λ_C is also a lognormal random variable with mean and standard deviation given by:

$$\begin{aligned} E\left[\lambda_C\right] &= k_0 \eta^{-k} \varepsilon_{IM} \exp\left(\frac{1}{2} k^2 \left(\sigma_{RC}^2 + \sigma_{UC}^2\right)\right) \\ Var\left[\ln \lambda_C\right] &= k^2 \sigma_{UC}^2 + \sigma_{UIM}^2 \end{aligned} \tag{A-24}$$

A.5. References

[1] Kennedy RP and Short SA. Basis for Seismic Provisions of DOE-STD-1020. Lawrence Livermore National Laboratory and Brookhaven National Laboratory, Washington, D.C., 1994.

[2] Cornell CA. Reliability-based earthquake-resistant design—the future, in *11th World Conference on Earthquake Engineering*, Acapulco, Mexico, 1996.

[3] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures. *Report No. RMS-35, RMS Program*,Stanford University, Stanford, CA, 1999. 357. http://www.stanford.edu/group/rms/

[4] Cornell CA, Jalayer F, Hamburger RO, and Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *Journal of Structural Engineering* 2002; **128**(4): 526–533.

[5] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1901-1914.

Appendix B: Effect of Predictability and Efficiency on Seismic Demand Hazard

This appendix illustrated the effect of predictability (i.e. aleatory uncertainty in a ground motion prediction equation) and efficiency (i.e. aleatory uncertainty in the seismic response, *EDP*|*IM* relationship). Only the derivation in terms of the rate-based formulation is given, as the formulation in the case of the probability-based formulation is significantly more involved.

B.1. Rate-based formulation

The rate-based formulation of the demand and intensity hazard, respectively are given by:

$$\lambda_{EDP}(edp) = \int_{IM} P_{EDP|IM}(edp|im)\left|\frac{d\lambda_{IM}(im)}{dIM}\right| dIM \tag{B-1}$$

$$\lambda_{IM}(im) = \sum_{i=1}^{Ns} \lambda_i \iint_{M,R} P_{IM|M,R,S}(im|m,r,s) f_{R|M}(r|m) f_M(m) dM dR \tag{B-2}$$

where $P_{EDP|IM}(edp|im)$ is the probability of $EDP > edp$ given $IM = im$; N_s is the number of seismic sources; λ_i is the 'activity rate' of source *i*; $P_{IM|M,R,S}(im|m,r,s)$ is the probability of $IM >$ im (obtained from a ground motion prediction equation); and $f_{R|M}(r|m)$ and $f_M(m)$ define the (conditional) distance and magnitude distributions of source *i*. Taking the derivative of Equation (B-2) noting that only $P_{IM|M,R,S}(im|m,r,s)$ depends on IM gives:

$$\frac{d\lambda_{IM}(im)}{dIM} = \sum_{i=1}^{Ns} \lambda_i \iint_{M,R} - f_{IM|M,R,S}(im|m,r,s) f_{R|M}(r|m) f_M(m) dM dR \tag{B-3}$$

where $f_{IM|M,R,S}(im|m,r,s)$ is the probability density function of the ground motion distribution and use has been made of the relation between the probability density and complementary cumulative distribution of a random variable [1]:

$$P_X(x) = \int_x^{\infty} f_X(x)dX \tag{B-4}$$

Further using Equation (B-4) for $P_{EDP|IM}(edp|im)$, Equation (B-1) becomes (following substitution of Equation (B-3)):

$$\lambda_{EDP}(edp) = \sum_{i=1}^{Ns} \lambda_i \iint_{M,Redp} \int_{IM}^{\infty} \left\{ \int_{IM} f_{EDP|IM}(edp|im) f_{IM|M,R,S}(im|m,r,s)dIM \right\} \mathrm{X} \\ dEDP f_{R|M}(r|m) f_M(m) dM dR \tag{B-5}$$

where all terms dependent on *IM* are contained within $\{\ \}$, thus giving the probability density function of *EDP* conditioned on *M,R,S*, specifically:

$$f_{EDP|M,R,S}(edp|m,r,s) = \int_{IM} f_{EDP|IM}(edp|im) f_{IM|M,R,S}(im|m,r,s)dIM \tag{B-6}$$

From Equation (B-6) it is clear that the choice of IM effects two distributions: $f_{EDP|IM}(edp|im)$, obtained from seismic response analysis and $f_{IM|M,R,S}(im|m,r,s)$ obtained from ground motion prediction equations. Generally, both $f_{EDP|IM}(edp|im)$ and $f_{IM|M,R,S}(im|m,r,s)$ are assumed to have lognormal distributions, and if it is assumed that the median of the EDP|IM distribution is given by aIM^b (i.e. $f_{EDP|IM}(edp|im) \sim LN(\ln[aIM^b], \sigma^2_{\ln EDP|IM})$ and $f_{IM|M,R,S}(im|m,r,s) \sim LN(\mu_{\ln IM|M,R,S}, \sigma^2_{\ln IM|M,R,S})$) then [1]:

$$f_{EDP|M,R,S}(edp|m,r,s) \sim LN(a\mu^b_{\ln IM|M,R,S}, \sigma^2_{\ln EDP|IM} + b^2\sigma^2_{\ln IM|M,R,S}) \tag{B-7}$$

That is, $f_{EDP|M,R,S}(edp|m,r,s)$ also has a lognormal distribution with the mean and variance specified. Hence the lognormal standard deviation of the *EDP|M,R,S* distribution is:

$$\sigma_{\ln EDP|M,R,S} = \sqrt{\sigma^2_{\ln EDP|IM} + b^2\sigma^2_{\ln IM|M,R,S}} \quad \text{(B-8)}$$

To further simplify Equation (B-8) Cornell *et al*. [2] note that $b = 1$ is a good approximation for the seismic response of moment frames.

B.2. References

[1] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[2] Cornell CA, Jalayer F, Hamburger RO, and Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. Journal of Structural Engineering 2002; 128(4): 526–533.

Appendix C: Derivation of the Statistical Moments for the Sum of the Product of Random Variables

Various integrals in the Pacific Earthquake Engineering Research (PEER) Centre's performance-based earthquake engineering (PBEE) framework [1] comprise of an integrand which is the product of two terms. An example is the annual frequency of structural collapse which is given by:

$$\lambda_C = \int_0^\infty P\left(C \middle| IM = im\right) \left| \frac{d\lambda(IM > im)}{dIM} \right| dIM \tag{C-1}$$

where λ_C = the annual rate of collapse; $P(C \mid IM = im)$ = the conditional probability of collapse given $IM = im$ (collapse fragility curve); and $\lambda(IM > im)$ = the annual rate of exceedance of $IM = im$ (ground motion hazard) at the site. The absolute value signs around the derivative of the ground motion hazard are used as its value is negative. For generality, Equation (C-1) and others can be expressed in the form:

$$Z = \int X(m) Y(m) dm \tag{C-2}$$

where the limits of integration are over the entire support of the variable m (typically $[0,\infty)$) .

While it is possible to obtain a closed-form solution to Equation (C-2) and other similar equations for specific functional forms of X and Y, their general solution requires the use of numerical integration (quadrature), or simulation (e.g. Monte Carlo) methods, with numerical integration methods more efficient than simulation methods for low-dimensionality integrals [2].

The numerical solution of Equation (C-2) involves the solution of the discretized function:

$$Z \approx \sum_{i=1}^{n} w_i X_i Y_i \tag{C-3}$$

where the domain of integration variable, m is discretized into n points, and w_i is an integration weight which is a function of the numerical integration scheme used.

Provided that the functions X and Y are such that the integrand XY is continuous and finite over the region of integration then the solution of Equation (C-2) (and the solution of Equation (C-3) as n tends to infinity) will be unique. However, if epistemic uncertainties are considered, such that the functions X and Y are themselves random variables, then Z will also a random variable with some distribution which we wish to determine. Again, there are various methods by which the distribution of Z can be determined. This appendix involves the derivation of the first two moments of the distribution of Z. The solution uses the properties of the expectation operator and is independent of the distributions of X and Y. It is important to note however, that the solution gives only the first two moments of the distribution of Z (higher moments can be computed, but are not done so here), and that some distribution must be fit to Z based on these moments.

C.1. The expectation of *Z*

From Equation (C-3), the expectation of Z can be computed from:

$$\begin{aligned} E[Z] &\approx E\left[\sum_{i=1}^{n} w_i X_i Y_i\right] \\ &= E\left[\{w_1 X_1 Y_1 + \ldots + w_n X_n Y_n\}\right] \\ &= \sum_{i=1}^{n} w_i E[X_i Y_i] \\ &= \sum_{i=1}^{n} w_i E[X_i] E[Y_i] \end{aligned} \tag{C-4}$$

where the final line is obtained based on the assumption that the X and Y are uncorrelated. Thus, the similarity in Equations (C-3) and (C-4) indicates that if epistemic uncertainties are ignored (although obviously they are always present) and the functions X and Y used are

assumed to be the 'best-estimate' (i.e. mean) values, then the value of *Z* obtained is in fact the mean value of *Z*.

C.2. The variance of *Z*

In order to compute the variance (i.e. second moment) of Z use is made of the definition of the variance:

$$\begin{aligned} Var[Z] &= E\left[(Z-\mu_Z)^2\right] \\ &= E\left[Z^2\right]-{\mu_Z}^2 \end{aligned} \tag{C-5}$$

where μ_Z is the mean of *Z*.

From Equation (C-3), the variance of *Z* can be computed from:

$$\begin{aligned} Var[Z] &\approx Var\left[\sum_{i=1}^{n} w_i X_i Y_i\right] \\ &= E\left[\left\{\sum_{i=1}^{n} w_i X_i Y_i\right\}^2\right]-E\left[\sum_{i=1}^{n} w_i X_i Y_i\right]^2 \\ &= E\left[\{w_1 X_1 Y_1+\ldots+w_n X_n Y_n\}^2\right]-\{E[w_1 X_1 Y_1]+\ldots+E[w_n X_n Y_n]\}^2 \\ &= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j X_i Y_i X_j Y_j\right]-\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j E[X_i Y_i]E[X_j Y_j] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j E[X_i X_j]E[Y_i Y_j]-\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j E[X_i]E[Y_i]E[X_j]E[Y_j] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \{E[X_i X_j]E[Y_i Y_j]-E[X_i]E[Y_i]E[X_j]E[Y_j]\} \end{aligned} \tag{C-6}$$

Where again the assumption that the *X* and *Y* are uncorrelated is used to obtain the fifth line in Equation (C-6). Letting $E[X_i]=\mu_{X_i}$ etc. and using the fact that $E[X_i X_j]=Cov[X_i X_j]+E[X_i]E[X_j]=\sigma_{X_i,X_j}+\mu_{X_i}$, Equation (C-6) can be simplified to:

$$\begin{aligned} Var[Z] &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \left\{\left(\sigma_{X_i,X_j}+\mu_{X_i}\mu_{X_j}\right)\left(\sigma_{Y_i,Y_j}+\mu_{Y_i}\mu_{Y_j}\right)-\mu_{X_i}\mu_{X_j}\mu_{Y_i}\mu_{Y_j}\right\} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \left\{\sigma_{X_i,X_j}\sigma_{Y_i,Y_j}+\mu_{X_i}\mu_{X_j}\sigma_{Y_i,Y_j}+\mu_{Y_i}\mu_{Y_j}\sigma_{X_i,X_j}+\mu_{X_i}\mu_{X_j}\mu_{Y_i}\mu_{Y_j}-\mu_{X_i}\mu_{X_j}\mu_{Y_i}\mu_{Y_j}\right\} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \left\{\sigma_{X_i,X_j}\sigma_{Y_i,Y_j}+\mu_{X_i}\mu_{X_j}\sigma_{Y_i,Y_j}+\mu_{Y_i}\mu_{Y_j}\sigma_{X_i,X_j}\right\} \end{aligned} \tag{C-7}$$

Thus the first and second moments of Z are given by:

$$E[Z] = \sum_{i=1}^{n} w_i \mu_{X_i} \mu_{Y_i} \tag{C-8}$$

$$Var[Z] = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \left\{ \sigma_{X_i,X_j} \sigma_{Y_i,Y_j} + \mu_{X_i} \mu_{X_j} \sigma_{Y_i,Y_j} + \mu_{Y_i} \mu_{Y_j} \sigma_{X_i,X_j} \right\} \tag{C-9}$$

C.3. References

[1] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).

[2] Kythe PK and Schaferhotter MR. *Handbook of Computational Methods for Integration*. Chapman and Hall/CRC, 2000;

Appendix D: Computation of Pseudo-Spectral Velocity, *PSV*, Correlation

In this appendix the mathematical details behind the computation of the correlation coefficient between the pseudo-spectral velocities of a ground motion at two different vibration periods as a function of the correlation between the pseudo-spectral acceleration at the same two periods.

D.1. Formulation

Given the relationship:

$$PSV_i = \frac{1}{\omega_i} Sa_i \tag{D-1}$$

where $Sa_i = Sa(T = T_i, 5\%)$ is the (pseudo) spectral acceleration; $PSV_i = PSV(T = T_i, 5\%)$ is the pseudo-spectral velocity, and the general expressions :

$$Cov[aX, bY] = abCov[X, Y] = ab\rho_{X,Y}\sigma_X\sigma_Y \tag{D-2}$$

$$Var[aX] = a^2 Var[X] = a^2\sigma_X^2 \tag{D-3}$$

Then it follows that:

$$\rho_{PSV_i,PSV_j} = \frac{Cov\left[\frac{1}{\omega_i}Sa_i, \frac{1}{\omega_j}Sa_j\right]}{\sqrt{Var\left[\frac{1}{\omega_i}Sa_i\right]Var\left[\frac{1}{\omega_j}Sa_j\right]}} = \frac{\frac{1}{\omega_i}\frac{1}{\omega_j}\rho_{Sa_i,Sa_j}\sigma_{Sa_i}\sigma_{Sa_j}}{\frac{1}{\omega_i}\sigma_{Sa_i}\frac{1}{\omega_j}\sigma_{Sa_j}} = \rho_{Sa_i,Sa_j} \quad \text{(D-4)}$$

Thus, the linearity of Equation (D-1) means that the correlation between spectral velocity terms is equal to that between spectral acceleration terms. As current correlation models for spectral accelerations use the natural logarithm of the spectral acceleration ordinate (i.e. ln*Sa*) then it is necessary to find a relationship between the correlation of the log and non-log forms of *Sa*.

The following expressions provide relationships between the moments (mean and variance) of the normal and lognormal univariate distributions [1] (where the substitutions $Y = lnSa$ and $X = Sa$ have been made for brevity):

$$\mu_X = \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right) \quad \text{(D-5)}$$

$$\sigma_X^2 = \exp\left(2\mu_Y + \sigma_Y^2\right)\left(\exp\left(\sigma_Y^2\right) - 1\right) \quad \text{(D-6)}$$

Equations (D-5) and (D-6) can be easily generalised for multivariate distributions to [2]:

$$\mu_{X_i} = \exp\left(\mu_{Y_i} + \frac{1}{2}\sigma_{Y_i}^2\right) \quad \text{(D-7)}$$

$$\sigma_{X_i}^2 = \exp\left(2\mu_{Y_i} + \sigma_{Y_i}^2\right)\left(\exp\left(\sigma_{Y_i}^2\right) - 1\right) \quad \text{(D-8)}$$

$$\sigma_{X_i,X_j} = \exp\left(\mu_{Y_i} + \mu_{Y_j} + \frac{1}{2}\left(\sigma_{Y_i}^2 + \sigma_{Y_j}^2\right)\right)\left(\exp\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right) - 1\right) \quad \text{(D-9)}$$

Making use of Equation (D-7), and that $\sigma_{X_i,X_j} = \rho_{X_i,X_j}\sigma_{X_i}\sigma_{X_j}$, Equation (D-9) can be rewritten as:

$$\rho_{X_i,X_j} = \frac{\mu_{X_i}}{\sigma_{X_i}} \frac{\mu_{X_j}}{\sigma_{X_j}} \left(\exp\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right)-1\right) \tag{D-10}$$

The first terms in Equation (D-9) are the reciprocal of the coefficient of variation of X_i and X_j, respectively, and can be expressed as a function of Y_i and Y_j using Equations (D-7) and (D-9):

$$\frac{\sigma_{Xi}}{\mu_{X_i}} = \frac{\sqrt{\exp\left(2\mu_{Y_i}+\sigma_{Y_i}^2\right)\left(\exp\left(\sigma_{Y_i}^2\right)-1\right)}}{\exp\left(\mu_{Y_i}+\frac{1}{2}\sigma_{Y_i}^2\right)} = \sqrt{\exp\left(\sigma_{Y_i}^2\right)-1} \tag{D-11}$$

Substituting Equation (D-11) into Equation (D-10) gives:

$$\rho_{X_i,X_j} = \frac{\exp\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right)-1}{\sqrt{\exp\left(\sigma_{Y_i}^2\right)-1}\sqrt{\exp\left(\sigma_{Y_j}^2\right)-1}} \tag{D-12}$$

Further, making the first order Taylor Series approximation $\exp(z) \approx 1+z+O\left(z^2\right)$ Equation (D-12) becomes:

$$\begin{aligned}
\rho_{X_i,X_j} &= \frac{\exp\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right)-1}{\sqrt{\exp\left(\sigma_{Y_i}^2\right)-1}\sqrt{\exp\left(\sigma_{Y_j}^2\right)-1}} \\
&= \frac{\left[1+\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}+O\left(\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right)^2\right)\right]-1}{\sqrt{\left[1+\sigma_{Y_i}^2+O\left(\sigma_{Y_i}^4\right)\right]-1}\sqrt{\left[1+\sigma_{Y_j}^2+O\left(\sigma_{Y_j}^4\right)\right]-1}} \\
&= \frac{\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}+O\left(\left(\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}\right)^2\right)}{\sqrt{\sigma_{Y_i}^2+O\left(\sigma_{Y_i}^4\right)}\sqrt{\sigma_{Y_j}^2+O\left(\sigma_{Y_j}^4\right)}} \\
&\approx \frac{\rho_{Y_i,Y_j}\sigma_{Y_i}\sigma_{Y_j}}{\sqrt{\sigma_{Y_i}^2}\sqrt{\sigma_{Y_j}^2}} \\
&= \rho_{Y_i,Y_j}
\end{aligned} \tag{D-13}$$

Thus by substituting $Y = lnSa$ and $X = Sa$:

$$\rho_{PSV_i,PSV_j} = \rho_{Sa_i,Sa_j} = \frac{\exp\left(\rho_{\ln Sa_i,\ln Sa_j}\sigma_{\ln Sa_i}\sigma_{\ln Sa_j}\right)-1}{\sqrt{\exp\left(\sigma_{\ln Sa_i}^2\right)-1}\sqrt{\exp\left(\sigma_{\ln Sa_j}^2\right)-1}} \approx \rho_{\ln Sa_i,\ln Sa_j} \tag{D-14}$$

D.2. References

[1] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.

[2] Johnson NL and Kotz S. *Distributions in Statistics: Continuous Multivariate Distributions* vol. 4. John Wiley & Sons, Inc., 1972;